

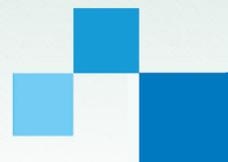


# nestor

Synergiepotenziale  
zwischen Grid- und  
eScience-Technologien  
für die digitale Langzeitarchivierung  
Udo Hönig und Wolfram Schiffmann

FernUniversität in Hagen  
Fakultät für Mathematik und Informatik - Lehrgebiet Rechnerarchitektur

nestor-materialien 12





Synergiepotenziale  
zwischen Grid- und  
eScience-Technologien  
für die digitale  
Langzeitarchivierung

Udo Hönig  
Wolfram Schiffmann

FernUniversität in Hagen  
Fakultät für Mathematik und Informatik  
Lehrgebiet Rechnerarchitektur

Herausgegeben von

**nestor** - Kompetenznetzwerk Langzeitarchivierung und  
Langzeitverfügbarkeit Digitaler Ressourcen für Deutschland

**nestor** - Network of Expertise in Long-Term Storage of Digital Resources

<http://www.langzeitarchivierung.de>

Projektpartner:

Bayerische Staatsbibliothek, München

Bundesarchiv

Deutsche Nationalbibliothek (Projektleitung)

FernUniversität in Hagen

Humboldt-Universität zu Berlin - Computer- und Medienservice / Universitätsbibliothek

Institut für Museumsforschung, Berlin

Niedersächsische Staats- und Universitätsbibliothek, Göttingen

© 2009

**nestor** - Kompetenznetzwerk Langzeitarchivierung und Langzeitverfügbarkeit  
Digitaler Ressourcen für Deutschland

Der Inhalt dieser Veröffentlichung darf vervielfältigt und verbreitet werden, sofern der Name des Rechteinhabers "nestor - Kompetenznetzwerk Langzeitarchivierung" genannt wird. Eine kommerzielle Nutzung ist nur mit Zustimmung des Rechteinhabers zulässig.

Betreuer dieser Veröffentlichung:

Niedersächsische Staats- und Universitätsbibliothek Göttingen

Heike Neuroth

Jens Ludwig

URN: <urn:nbn:de:0008-2009012115>

<http://nbn-resolving.de/urn:nbn:de:0008-2009012115>

## Grid-Technologie und Langzeitarchivierung in nestor

Die modernen Informationstechnologien haben in allen Lebensbereichen starke Veränderungen bewirkt. Besonders stark beeinflusst sind die Wissenschaften, die auch eine treibende Kraft dieser Entwicklungen sind und immer größere Anforderungen an Rechner, Speicher und IT-Werkzeuge stellen. In neuen Experimenten der Teilchenphysik werden kaum bewältigbare Datenmengen für Tausende von Wissenschaftlern produziert, Klimaforscher berechnen immer detailliertere Modelle des Systems Erde, und die Geisteswissenschaften beginnen riesige digitale Sammlungen von Kulturgütern mit Rechnern zu analysieren. Die Grid-Technologie zur Aufteilung der Aufgaben auf viele verteilte IT-Ressourcen ist ein Mittel, um den Herausforderungen dieser neuen, als e-Science bezeichneten wissenschaftlichen Arbeitsweise gerecht zu werden.

nestor und Wissenschaftler weltweit haben immer wieder darauf hingewiesen, dass mit der Zunahme der Bedeutung digitaler Daten auch die Notwendigkeit wächst, ihre langfristige Nutzbarkeit zu sichern. Bei der Grid-Technologie ergibt sich die chancenreiche Situation, dass nicht nur wertvolle und zu erhaltende Daten produziert werden, sondern auch Mittel bereit gestellt werden, die für die Herausforderung der Langzeitarchivierung großer und komplexer Datenmengen nutzbar sein können. Die klassischen Gedächtnisorganisationen - wie Bibliotheken, Archive und Museen - und die neuen Gedächtnisorganisationen - wie Daten- und Rechenzentren - können wechselseitig voneinander profitieren.

Um dieses Potenzial auszuloten, hat nestor in seiner zweiten Projektphase eine Arbeitsgruppe mit Fachleuten aus klassischen Gedächtnisinstitutionen und aus e-Science- und grid-engagierten Institutionen initiiert und drei Expertisen in Auftrag gegeben. Diese Expertisen untersuchen den Ist-Stand und die Anforderungen und Ziele für das Zusammenspiel von e-Science-/Grid-Technologie und Langzeitarchivierung unter drei Gesichtspunkten:

Welche Anforderungen gibt es für die Archivierung von Forschungsdaten?

Was sind die möglichen Synergien, die angestrebt werden sollten?

Und auf welche Standards können weitere Arbeiten in diesen Bereich aufgebaut werden und welche sind gegebenenfalls noch zu entwickeln?

Neben der Untersuchung des Standes der Technik, sind einige Projekte der deutschen Grid-Initiative D-Grid befragt worden. nestor wird in seiner Grid-/eScience-Arbeitsgruppe die Ergebnisse der Expertisen aufnehmen und versuchen, eine Landkarte für die weiteren Entwicklungsperspektiven zu zeichnen.

e-Science-/Grid-Technologie und Langzeitarchivierung sind relativ neue Forschungsbereiche, die sich sehr schnell entwickeln. Einzelne Fragen, die von nestor Mitte 2006 formuliert wurden, als die ersten Projekte der deutschen Grid-Initiative D-Grid gerade gestartet waren, stellen sich heute, wo bald schon die dritte Generation von D-Grid-Projekten beginnt, unter den veränderten Bedingungen möglicherweise anders dar.

Die Expertisen müssen daher auch vor ihrem Entstehungshintergrund betrachtet werden. Derzeit liefern sie eine Beschreibung sinnvoller und notwendiger Entwicklungen. Wenn sie in naher Zukunft „veralten“, weil sie zur erfolgreichen Zusammenarbeit von e-Science/Grid und Langzeitarchivierung beigetragen haben, dann haben sie ihren Sinn erfüllt.

**Stellungnahme zur Expertise  
"Synergiepotenziale zwischen Grid- und  
eScience-Technologien für die digitale Langzeitarchivierung"**

**Tobias Blanke, King's College London, 1.9.2008**

Die Expertise von Udo Hönig und Wolfram Schiffmann zu „Synergiepotenzialen zwischen Grid- und eScience-Technologien für die digitale Langzeitarchivierung (LZA)“ gibt eine interessante Perspektive auf das gegenwärtig intensiv diskutierte Thema einer Integration von Archiven mit eScience-Umgebungen. Im Weiteren werden zunächst Resultate der Expertise diskutiert, bevor am Ende weitere vertiefende Analysen vorgeschlagen werden.

Die Expertise führt zunächst in einige der vorhandenen Technologien ein. Detailliert wird die Gruppe der fortschrittlichen Netzwerktechnologien, die unter dem Namen Grid zusammengefasst sind, auf ihre Nutzbarkeit in einem LZA-Zusammenhang untersucht. Die Forschung hat seit langem vor allem im Bereich der so genannten Datengrids offensichtliche Synergieeffekte identifiziert. So wird z.B. der Storage Resource Broker (SRB) des San Diego Super Computer Centre schon von vielen Archivorganisationen in der Produktion benutzt. SRB ist im Wesentlichen ein einheitliches Interface zu heterogenen Datenressourcen mit vielen zusätzlichen Diensten wie automatische Replikation und Datensicherung. Existierende Projekte zur flexiblen Integration von Datengrids mit LZA-Systemen werden zwar nicht im Einzelnen in der Expertise ausgeführt, sind aber sicherlich dennoch Grundlage der Argumentation. Als Beispiele sind hier die Australische DART Initiative zu nennen, die eine Standardimplementation der Integration von SRB und Fedora realisiert hat,<sup>1</sup> oder das Projekt zur Integration von iRODS und Fedora.<sup>2</sup>

Im letzten Kapitel der Expertise werden schließlich die Synergiepotenziale von Grid und LZA-Systemen in einer Taxonomie zusammengefasst. Aus dieser wird eine service-orientierte integrierte Grid-LZA-Infrastruktur hergeleitet. Solche service-oriented architectures (SOA) werden nicht nur von den Autoren als Erfolg versprechender Mittelweg angesehen. SOA's bieten Interfaces, die es ermöglichen, existierende Systeme ohne größere Neuimplementierung miteinander zu verbinden. So werden die langjährigen Erfahrungen, die in Referenzimplementierungen von LZA-Systemen eingeflossen sind, nicht einfach aufgegeben zugunsten einer kompletten Neuimplementierung. Andere Referenzimplementierungen wie Fedora sind schon als SOA ausgelegt. Auf der anderen Seite können Grid-Dienste bestehende Funktionalitäten von SOA's verbessern. Die Übertragung von Daten ist z.B. effizienter mit GridFTP möglich als mit vergleichbaren Standards, die in SOA-Implementierungen üblich sind. Die Autoren raten schließlich von einer vollkommenen Gridifizierung der LZA-Systeme ab. Das könnte zu Sicherheitsrisiken und Systeminstabilitäten führen.

Um den Bedarf an LZA-Systemen in eScience besser zu verstehen, haben die Autoren drei D-Grid Projekte kontaktiert: AstroGrid, C3-Grid (Community Climate Collaboration Grid) und TextGrid. Alle drei planen oder visieren zumindest an, größere Datenarchive in ihre Umgebungen einzubinden. In der gegenwärtigen deutschen eScience-Initiative setzt sich das Bewusstsein mehr und mehr durch, dass Daten nicht nur für den Gebrauch gegenwärtiger wissenschaftlicher Analyse bereitzustellen sind. Vielmehr ist es notwendig, Daten über einen längeren Zeitraum zu erhalten, damit auf ihnen basierende wissenschaftliche Resultate

---

1 <http://www.itee.uq.edu.au/~eresearch/projects/dart/outcomes/FedoraDB.php>

2 <https://www.irods.org/index.php/Fedora>

effektiv und vor allem langfristig überprüft werden können. Aus den Interviews wird deutlich, dass es zwar Gemeinsamkeiten zwischen den verschiedenen Disziplinen gibt, aber dennoch weitgehend unabhängige, disziplinspezifische Lösungen vorherrschen. Es gibt offensichtlich große Unterschiede, was die einzelnen Forschungsbereiche als relevantes und erhaltenswertes Material ansehen. TextGrid bedient eine vorwiegend geisteswissenschaftliche Gemeinde, die auch auf Jahrzehnte nach der Erstellung einer Digitalisierung eines Artefakts noch Zugriff auf diese erwartet. TextGrid arbeitet damit in einer ähnlichen Zeitperspektive wie Bibliotheken oder größere Archive. In den mehr naturwissenschaftlich orientierten Disziplinen scheint man mit einer wesentlich kürzeren Archivierung zufrieden. Jenseits von 10 Jahren scheint die Bedeutung des Forschungsmaterials abzunehmen.

Solche Unterschiede in der Wahrnehmung und Implementierung von LZA-Systemen setzen sich fort in Schwierigkeiten, dauerhaft interdisziplinäre Kollaboration zu bewerkstelligen. Diese gibt es nicht nur in Deutschland lediglich in Ansätzen, obgleich gerade von einem solchen Austausch große Synergiepotenziale zu erwarten sind. Es ist zwar nicht anzunehmen, dass z.B. Metadatenformate jemals gänzlich gemeinsam genutzt werden können, aber eines der Versprechen des eScience-Programms ist der wissenschaftliche Austausch über Disziplinengrenzen hinweg. Ein Archäologe z.B. kann sehr an der Auswertung von astronomischen Messungen der Erdoberfläche interessiert sein, um seine Ausgrabungen zu unterstützen. Insofern wäre es wünschenswert, könnte man direkt aus der geisteswissenschaftlichen Community-Umgebung auf Archivdaten der Astronomen zugreifen. Ein solches Programm, das Interoperabilitätsinitiativen von Archivsystemen wie JISC CRIG3 mit vergleichbaren Initiativen in Grid-Systemen zusammenbrächte, führt jedoch über den Umfang der vorliegenden Expertise hinaus. Es leitet über, zu möglichen weiteren Schritten in der Analyse der potenziellen Synergie von eScience-Technologien und LZA-Systemen, die kurz angerissen werden sollen.

Insgesamt gibt die Expertise von Udo Hönig und Wolfram Schiffmann einen guten Überblick über Synergiepotenziale von Grids und digitalen LZA-Systemen. Zukünftige Studien könnten an einer weiterführenden, detaillierteren Perspektive arbeiten, die die Nützlichkeit von LZA-Systemen für eScience-Umgebungen ausführten. Aus der Perspektive Großbritanniens und sicherlich auch Europas werden viele eScience-Entwicklungen immer noch wesentlich von der Hochenergiephysik (HEP) getragen. Früher wurden die Daten eines abgelaufenen HEP-Experiments einfach verworfen. Mit den immensen Kosten für neuere Experimente wie dem Large Hadron Collider (LHC) in CERN hat auch hier ein Umdenken eingesetzt.<sup>4</sup> Experimente können nicht mehr einfach wiederholt werden. Daher wird auch in diesem Bereich mit seinen ungeheuren Datenmengen in der Zukunft LZA eine größere Rolle spielen. Da es jedoch im Augenblick unrealistisch ist, die gesamte Versuchsanordnung solcher Experimente zusammen mit den unglaublich großen Datenmengen abzuspeichern, ist großer Forschungsbedarf vorhanden, damit LZA für HEP nützlich wird. Es wäre auch interessant gewesen, in die Expertise eine Industrieperspektive einzubinden. Grids werden mittlerweile auch in stärker kommerziell ausgerichteten Einrichtungen stark genutzt. Dass es ein kommerzielles Interesse an z.B. Datengrids gibt, hat ja gerade die Entwicklung von Nirvana angestoßen. Man würde vermuten, dass Technologien zur LZA und Grids in der Industrie eher für hausinterne Lösungen bereitgestellt werden. Der Vorschlag zu weiteren Studien zu HEP und Industrie sind als Anregungen zu verstehen, weitere, detaillierte Studien zu erarbeiten, die auf der vorliegenden Expertise aufbauen könnten.

---

3 <http://www.ukoln.ac.uk/repositories/digirep/index/CRIG>

4 <http://www.computerweekly.com/Articles/2008/08/06/231762/in-search-of-the-big-bang.htm>



**Expertise**  
**“Synergiepotenziale zwischen Grid- und  
eScience-Technologien für die digitale Langzeit-  
archivierung“**

---

FernUniversität in Hagen  
Fakultät für Mathematik und Informatik  
Lehrgebiet Rechnerarchitektur

Udo Hönig  
Wolfram Schiffmann



## Inhaltsverzeichnis

Zusammenfassung .....	4
1 Motivation und Ziele .....	5
2 Langzeitarchivierung.....	6
2.1 OAIS-Referenzmodell.....	7
2.1.1 Ingest .....	9
2.1.2 Archival Storage.....	9
2.1.3 Data Management .....	9
2.1.4 Administration.....	10
2.1.5 Preservation Planning.....	10
2.1.6 Access.....	10
2.2 Beispiele für OAIS-kompatible Archive.....	10
2.2.1 BABS .....	10
2.2.2 kopal .....	11
2.2.3 Institutionelle Repositorien .....	11
3 Grid Computing und eScience .....	12
3.1 Grid Computing.....	15
3.1.1 Rechen-Grids.....	15
3.1.2 Daten-Grids .....	20
3.2 eScience.....	24
3.2.1 eSciDoc .....	25
3.2.2 ONTOVERSE .....	26
3.2.3 WIKINGER.....	26
3.2.4 Im Wissensnetz .....	27
4 Befragung von drei CG-Projekten.....	28
4.1 AstroGrid-D .....	28
4.2 C3-Grid.....	29
4.3 TextGrid .....	32
4.4 Zusammenfassung .....	33
5 Synergiepotenziale .....	36
5.1 Ingest-Prozess.....	39
5.2 Access-Prozess .....	40
5.3 Preservation-Prozess .....	40
5.4 Archival Storage.....	41
6 Handlungsempfehlungen.....	42
7 Literatur .....	45
8 Anhang .....	48
8.1 Deutsche Projekte.....	48
8.2 Europäische Projekte.....	49
8.3 Internationale Projekte .....	50

## Zusammenfassung

Die vorliegende Expertise soll mögliche Synergiepotenziale zwischen Grid- und eScience-Technologien für die digitale Langzeitarchivierung (dLZA) aufzeigen und Ansatzpunkte für Integrationskonzepte entwickeln. Die Expertise beschränkt sich primär auf den technischen Fokus, d.h. es werden keine organisatorisch-strategischen Aspekte behandelt.

Zunächst wird in die Grundlagen der dLZA eingeführt und das Open Archival Information System (OAIS) vorgestellt, das eine technologieunabhängige Basisarchitektur für die LZA spezifiziert. Anhand von Beispielen wird gezeigt, wie diese Architektur sowohl von großen Bibliotheken (so genannte Gedächtnisorganisationen) als auch von kleineren Instituten (meist Universitäten) zur Erhaltung und Bereitstellung von Informationen oder Inhalten eingesetzt wird.

Mit Hilfe von Grid-Technologien kann man geographisch verteilte, heterogene Ressourcen nutzen, die nicht unter einer gemeinsamen Administration stehen. Bei den Ressourcen kann es sich um Computer, Speicher oder aber auch fernsteuerbare Geräte (z.B. Messgeräte) handeln. Ein Grid entsteht, indem man die Ressourcen mit Hilfe einer Middleware über ein Netzwerk koppelt und sie so zu einer scheinbar einzigen, aber extrem leistungsfähigen Infrastruktur vereinigt. Im Wesentlichen kann man Rechen- und Daten-Grids unterscheiden, die für so genannte virtuelle Organisationen über das Internet nutzbare Dienste in Form von Rechenleistung oder Speicherplatz anbieten.

Unter eScience versteht man sowohl die orts- und/oder zeitunabhängige Zusammenarbeit in Kerngebieten der Wissenschaft als auch die dazu benötigten Werkzeuge und Infrastrukturen. Dabei bilden die o.g. Grid-Technologien die Basis-Infrastruktur für eScience und viele fachspezifische Arbeitsgruppen betreiben ihre eigenen Community Grids (CGs), die an die besonderen Anforderungen des jeweiligen Fachgebiets angepasst sind.

In der vorliegenden Expertise wurden exemplarisch drei CGs und das Word Data Center Climate (WDCC) befragt, um einerseits den aktuellen Entwicklungsstand ihrer eScience-Umgebungen und andererseits ihre Anforderungen an eine Grid-basierte dLZA-Infrastruktur zu analysieren. Mögliche Synergiepotenziale werden anhand der Archiv-Prozesse des OAIS-Referenzmodells identifiziert. Es wird eine serviceorientierte LZA-Infrastruktur vorgeschlagen, die auf vorhandenen Grid-Technologien aufbaut und es gleichzeitig ermöglicht, sowohl existierende als auch neue Archive nahtlos in eine Grid-Umgebung zu integrieren.

# 1 Motivation und Ziele

Trotz vieler Vorteile elektronischer Dokumente wirft ihre immer stärkere Verbreitung ein großes Problem für die Zukunft auf. Digital gespeicherte Inhalte sind nämlich sehr viel schwieriger zu erhalten als Bücher oder andere an physische Medien gebundene Dokumente. Gleichzeitig steigt das zu erhaltende Datenvolumen explosionsartig an. Dies ist damit zu erklären, dass durch Simulationen und rechnergestützte Beobachtungen immer mehr Daten pro Zeiteinheit erzeugt werden. So produzieren z.B. die Experimente aus dem Bereich der Hochenergie-Physik pro Jahr ca. 15 PetaBytes an Daten (1 PetaByte entspricht 1 Million GByte). Die erforderlichen Speichersysteme müssen also gigantische Kapazitäten bereitstellen. Man schätzt, dass auf der Welt z.Zt. ein Datenvolumen von 161 Exabyte (161 Milliarden GByte) vorliegt und dass alle 15 Minuten etwa soviel Information hinzukommt, wie momentan schon in der amerikanischen Kongress-Bibliothek vorhanden ist. Natürlich ist es weder möglich noch sinnvoll, diese Menge an Information vollständig und für alle Zeiten zu bewahren. Es muss also sorgfältig ausgewählt werden, welche Informationen langfristig gespeichert werden.

Für die Langzeitarchivierung im Allgemeinen (LZA) sind ebenso wie für die digitale Langzeitarchivierung (dLZA) im Besonderen so genannte Gedächtnisorganisationen (wie bspw. große Bibliotheken) zuständig. Für sie stellen die gigantischen Datenmengen und sich immer steigenden Zuwachsraten eine große Herausforderung dar. Gedächtnisorganisationen betreiben Archive, die über Speichersysteme mit enormen Speicherkapazitäten verfügen müssen. Um die eingelagerten digitalen Inhalte über Jahrzehnte bzw. –hunderte zu erhalten, müssen die Archive entsprechende Verwaltungs- und Organisationsstrukturen aufbauen. Da sich die logische Repräsentation von Dokumenten im Laufe der Zeit immer wieder ändert, müssen in bestimmten Zeitabständen auch rechenintensive Formatumwandlungen durchgeführt werden. Viele der neu hinzukommenden Inhalte werden heute in Grid- bzw. eScience-Umgebungen bereits digital „geboren“. Es ist daher naheliegend, die dort verwendeten Techniken und Ressourcen auch für die dLZA einzusetzen bzw. eine enge Verzahnung dieser Bereiche herzustellen. Ebenso können in umgekehrter Richtung Mehrwerte durch das Vorhandensein von dLZA-Funktionalitäten entstehen. Mit dem aus dem griechischen stammenden Begriff *Synergie* bezeichnet man im Allgemeinen „das Zusammenwirken verschiedener Kräfte, Stoffe oder Faktoren derart, dass die Gesamtwirkung größer ist als die Summe der Wirkungen der Einzelkomponenten“ [Meyers Lexikonverlag, 2007]. Den Effekt der gegenseitigen Verstärkung durch Zusammenwirken bezeichnet man als *Synergieeffekt*.

Die vorliegende Expertise soll mögliche Synergiepotenziale zwischen Grid- und eScience-Technologien für die digitale Langzeitarchivierung aufzeigen und Ansatzpunkte für Integrationskonzepte entwickeln. Die Expertise beschränkt sich primär auf den technischen Fokus, d.h. es werden keine organisatorisch-strategischen Aspekte behandelt. Außerdem betrachtet die vorliegende Expertise ausschließlich elektronisch gespeicherte Daten und Dokumente, weshalb nachfolgend ohne Verlust der Eindeutigkeit vereinfachend von „Langzeitarchivierung“ bzw. „LZA“ gesprochen werden kann. Zunächst wird auf Grundlagen und Methoden der LZA eingegangen, bevor anhand einer Auswahl relevanter Projekte der aktuelle Stand in den Bereichen Grid Computing und eScience zusammengefasst wird. Im zweiten Teil der Expertise werden die benötigten Eigenschaften einer zukünftigen Infrastruktur mit Synergiepotenzialen für die LZA beschrieben und es werden Handlungsempfehlungen gegeben, wie

eine solche Infrastruktur unter Einbeziehung bereits vorhandener Projektentwicklungen realisiert werden kann.

## 2 Langzeitarchivierung

Wie bereits im vorangegangenen Kapitel erläutert wurde, konzentriert sich die vorliegende Expertise ausschließlich auf die im Kontext der Langzeitarchivierung digitaler Daten relevanten Aspekte. Gegenstand der LZA sind demnach Informationen in digitaler Darstellung, die mit Speichermedien mehr oder weniger lange gespeichert werden können. Die am häufigsten eingesetzten Speichermedien sind Festplatten, CD/DVDs und Magnetbänder, wobei die angegebene Reihung den erreichbaren Zugriffsgeschwindigkeiten entspricht. Da die maximalen Speicherkapazitäten eines einzelnen Speichermediums relativ gering sind (Festplatten ca. 500 GBytes, DVD ca. 5 GBytes, Magnetband ca. 20 GBytes), müssen zur LZA mehrere solcher Speichermedien parallel betrieben werden. Die technischen Grundlagen von Speichermedien sind im Rahmen dieser Expertise nur von untergeordneter Bedeutung. Sie können in einschlägigen Lehrbüchern nachgelesen werden [Schiffmann, 2005].

Die zu erhaltenden Informationen oder *Inhalte* können in drei Kategorien eingeteilt werden: 1. Reine Daten (auch Rohdaten), 2. Formatierte Dokumente (zu denen wir im weiteren Verlauf dieser Studie auch unterschiedliche Kategorien statischer und dynamischer medialer und multimedialer digitaler Objekte wie z.B. Audio- und Videodokumente zählen) und 3. Software.

Rohdaten können z.B. aus Messwerten von einem Experiment stammen. Die Messwerte sind nur dann brauchbar, wenn zusätzliche Information über den Dateiaufbau gegeben werden. Formatierte Dokumente sind entweder Veröffentlichungen (z.B. im pdf-Format) oder Audio- bzw. Video-Dateien. Die Inhalte dieser Dokumente sind unabhängig von Benutzereingaben oder anderen Laufzeiteinflüssen, d.h. sie können nur angeschaut bzw. angehört werden, wie sie sind. Es handelt sich somit um statische Dokumente. Im Gegensatz dazu gibt es auch dynamische Dokumente, deren Erscheinungsbild durch Benutzereingaben interaktiv gesteuert werden kann. Beispiele dafür sind ein mit einem Autorensystem erstelltes Dokument, wie z.B. ein Computer Based Training (CBT) oder ein interaktives Simulationsergebnis. Solche dynamischen Dokumente verhalten sich ähnlich wie Software. Um sie langfristig zu bewahren, ist es in der Regel nötig, auch die dazu passende Ablaufumgebung zu archivieren.

Eine Gedächtnisorganisation muss im Wesentlichen zwei Aufgaben erfüllen: Sie muss einerseits ein Bitmuster langfristig (über ein oder mehrere Jahrzehnte) speichern und andererseits die Interpretierbarkeit der gespeicherten Bitmuster sicherstellen. Alle Speichermedien haben nur begrenzte Haltbarkeit, d.h. die Bitmuster müssen in regelmäßigen Zeitabständen von den Speichermedien gelesen und auf andere Speichermedien kopiert werden (Datenträgermigration). Zur Fehlertoleranz gegenüber Schreib- und Lesefehlern verwenden Speichermedien Prüfinformationen und fehlerkorrigierende Codes. Anhand der Prüfinformation kann die Integrität der gespeicherten Daten überprüft und mit Hilfe fehlerkorrigierender Codes eventuell auch wiederhergestellt werden. Trotzdem kann auf einem einzigen Speichermedium keine völlig sichere Speicherung eines Bitmusters garantiert werden. Daher speichert man ein zu erhaltendes Bitmuster redundant auf zwei oder mehr Speichermedien, um Ausfälle zu tolerieren und so die Speichersicherheit zu erhöhen. Der Grad der Fehlertoleranz ist proportional

zur Zahl der verwendeten Repliken. Andererseits steigt jedoch auch der Aufwand, d.h. es muss ein Kompromiss zwischen der Anzahl der Repliken und dem Grad der Fehlertoleranz gefunden werden.

Neben der Erhaltung der Bitmuster eines Archivierungsobjekts muss dessen Interpretierbarkeit für den Menschen bewahrt werden. Bei Rohdaten muss sowohl die Anordnung der Messwerte als auch deren Codierung (z.B. als Gleitkommazahlen im IEEE-754-Standard) bekannt sein. Ähnlich wird bei statischen Dokumenten das Dateiformat von einem Anzeigeprogramm gelesen, interpretiert und für die menschlichen Sinne aufbereitet. Im Laufe der Zeit ändern sich jedoch die Standards für Dokumente. Es wird daher früher oder später vorkommen, dass auf nachfolgenden Systemplattformen (Computer-Hardware und -Betriebssystem) die alten Dateiformate nicht mehr durch entsprechende Anzeigeprogramme unterstützt werden.

Um die Lesbarkeit eines Dokuments zu erhalten, kann die Datei in das neue Format konvertiert und dann von dem aktuellen Anzeigeprogramm dargestellt werden. Der Nachteil dieses als *Migration* bezeichneten Verfahrens ist, dass die so genannte *Authentizität* des Dokuments gefährdet ist, wenn bei einer Konvertierung ein Fehler auftritt, der erst später bemerkt wird. Um dieses Problem zu lösen, sollte man stets das Original-Dokument aufbewahren und sämtliche Veränderungen daran protokollieren. Eine Alternative zur Migration stellt die so genannte *Emulation* dar. Hierbei wird die zum ursprünglichen Anzeigeprogramm passende Rechnerplattform erhalten, indem sie auf einer aktuellen Rechnerplattform nachgebildet (emuliert) wird. Die Emulation entspricht also der Migration der Rechnerplattform und ermöglicht es, sowohl dynamische Dokumente als auch Software langfristig zu archivieren. Da hierbei das Original-Dokument zusammen mit dem (Original-)Anzeigeprogramm erhalten bleibt, kann die Authentizität einfacher sichergestellt werden. Darüberhinaus entfällt auch der nicht unerhebliche Rechenaufwand, um bei umfangreichen Archiven die Dateiformate aller betroffenen Dokumente zu konvertieren.

## **2.1 OAIS-Referenzmodell**

Das Open Archival Information System (OAIS) definiert ein Archiv als eine Organisation, in der Menschen mit technischen Systemen zusammenwirken, um Informationen über lange Zeiträume zu erhalten. Gleichzeitig soll es berechtigten Benutzern ermöglichen, auf die archivierten Inhalte zuzugreifen. Da die Erhaltungszeiträume in der Größenordnung von mehreren Jahrzehnten oder gar Jahrhunderten liegen, ist es sehr wichtig, dass das OAIS-Referenzmodell nur die zugrunde liegende Architektur spezifiziert und somit von einem konkreten Aufbau und einer technischen Implementierung abstrahiert. Das OAIS-Referenzmodell ist ein Standard, der vom Consultative Committee für Space Data Systems (CCSDS) im Auftrag der International Organization for Standardization (ISO) entwickelt wurde und nach achtjährigen Forschungsarbeiten im Februar 2003 als ISO-Norm veröffentlicht wurde. Die Grundzüge dieses sehr detaillierten Modells (vgl. [OAIS, 2002]) sollen im Folgenden kurz beschrieben werden, um dann später mögliche Synergiepotenziale zu identifizieren. Das OAIS-Referenzmodell ist zwar ein allgemeines Modell für die Organisation eines Archivs, das sowohl physische als auch digitale Bestände haben kann. Heutzutage steht jedoch die *digitale LZA* im Vordergrund.

Ein wesentliches Merkmal des OAIS-Referenzmodells ist die Unterscheidung zwischen Daten und Information. Daten, oft auch Bitströme genannt, sind nur die Träger der Information. Auf den Speichermedien von Computersystemen können daher nur *Daten*, nicht aber *Informationen* für einen begrenzten Zeitraum (typ. bis max. 10 Jahre) abgelegt werden. *Information* entsteht dagegen erst dann, wenn die Daten mit Hilfe von Darstellungsinformationen (Repräsentations-Information) so umgeformt werden, dass sie eine Bedeutung erhalten. So entspricht zum Beispiel dem Bitstrom 1110 dem dezimalen Zahlenwert 14, sofern wir das duale Zahlensystem zugrunde legen. Bei oktaler Zahlendarstellung würde sich dagegen die Ziffernfolge  $16_{\text{okt}}$  ergeben. Um Informationen richtig interpretieren zu können, muss beim Empfänger also auch eine entsprechende Wissensbasis (Knowledge Base) vorhanden sein. Die Menschen verstehen intuitiv Darstellungen im Dezimalsystem. Man benutzt daher das Dezimalsystem als Standard-Wissensbasis bei der Interpretation von Zahlen und gibt nur bei anderen Zahlensystemen zusätzliche Informationen über das verwendete Zahlensystem, wie gezeigt, als Index an.

Aus den obigen Überlegungen folgt, dass wir zur Erhaltung von Informationen nicht nur Daten, sondern auch die zugehörigen Darstellungsinformationen benötigen. Beides zusammen bezeichnet man als Inhaltsinformationen oder Content Information (CI). Diese wiederum werden ergänzt um Zusatzinformationen, die z.B. zur Bewahrung der Integrität der Inhaltsinformationen dienen und die als Preservation Description Information (PDI) bezeichnet werden. Zusammen bilden CI und PDI das Information Package (IP), das die elementare Informationseinheit eines Archivs darstellt.

Man unterscheidet im OAIS-Referenzmodell drei Arten von IPs:

1. SIP steht für Submission Information Package. Es wird von einem Erzeuger (Producer) zur Archivierung eingereicht.
2. AIP steht für Archival Information Package. Es wird innerhalb des Archivs zur eigentlichen Langzeit-Speicherung (Archival Storage) verwendet.
3. DIP steht für Dissemination Information Package. Es bezeichnet die Daten, die der Verbraucher (Consumer) vom Archiv erhält.

Um die einzelnen IPs leichter verwalten und wiederfinden zu können, werden zusätzliche Beschreibungen erzeugt, die kompaktere und trotzdem aussagekräftige Informationen über den Inhalt der IPs enthalten. Diese Metadaten (Descriptive Information) sind zur Verwaltung der eingelagerten AIPs und zur Beschleunigung der Abläufe innerhalb eines Archivs von großer Bedeutung. Das OAIS-Referenzmodell spezifiziert detaillierte Strukturen für den Informationsfluss und die benötigten Funktionen innerhalb eines Archivs. Eine ausführliche Darstellung würde den Rahmen dieser Expertise sprengen (siehe [OAIS, 2002]). Im Folgenden werden lediglich die sechs grundlegenden Einheiten und eine vereinfachte Systemstruktur gemäß Abbildung 1 beschrieben.



## **2.1.4 Administration**

Die Administration-Komponente kontrolliert die Gesamtfunktion des Archivs. Hierzu zählen die Konfiguration der Hard- und Software sowie deren Optimierung. Außerdem werden durch die Administration auch die Standards für die zulässigen SIPs festgelegt und überprüft, ob die von den Erzeugern eingereichten Inhalte diese Anforderungen erfüllen. Ebenso entscheidet die Administration über die Archiv-internen Standards und veranlasst bei Änderungen dieser Standards die Formatmigration bei den AIPs. Daneben sind von der Administration auch viele organisatorische Aufgaben zu lösen, die aber im Rahmen der vorliegenden Expertise nicht näher behandelt werden.

## **2.1.5 Preservation Planning**

Das Preservation Planning soll sicherstellen, dass die gespeicherten Inhalte auch dann noch verfügbar sind, wenn die technische Infrastruktur veraltet bzw. gar nicht mehr vorhanden ist. Hierzu müssen ähnliche Aufgaben wie bei der Administration gelöst werden. Allerdings wird hier ein größerer Zeithorizont zugrunde gelegt und die gesamte Hard- und Softwareumgebung des Archivs einbezogen. Zum Preservation Planning zählen auch die Planungen von umfangreichen Migrationsläufen sowie die Entwicklung und der Test neuer Softwarekomponenten.

## **2.1.6 Access**

Wenn ein Verbraucher (Consumer) auf das Archiv zugreifen möchte, muss zunächst geprüft werden, ob er dazu berechtigt ist. Dann benötigt er Dienste, um gezielt nach Inhalten zu suchen. Neben diesen Funktionen koordiniert die Access-Komponente auch die Zugriffe auf die Inhalte, indem sie aus AIPs die DIPs aufbereitet und schließlich an den Benutzer ausliefert.

## **2.2 Beispiele für OAIS-kompatible Archive**

Der OAIS-Standard bildet die Grundlage für viele Archive. Als Beispiele sollen hier kurz BABS, kopal und institutionelle Repositorien beschrieben werden. Während bei den beiden erstgenannten die Langzeitarchivierung im Vordergrund steht, sind die Erhaltungszeiträume und die Erhaltungssicherheit bei den institutionellen Archiven deutlich geringer.

### **2.2.1 BABS**

Das Projekt Bibliothekarisches Archivierungs- und Bereitstellungssystem (BABS) wird seit 2005 von der DFG gefördert und basiert auf einer Kooperation der Bayrischen Staatsbibliothek (BSB) und des Leibniz-Rechenzentrums (LRZ). Das Ziel des auf zwei Jahre ausgelegten Projekts besteht darin, eine organisatorische und technische Infrastruktur zur Langzeitarchivierung zu schaffen. BABS dient als exemplarische Implementierung, um Erfahrungen mit Verfahren des Datenmanagements und neuen Workflows zu sammeln. Die Architektur von

BABS basiert auf dem oben beschriebenen OAIS-Referenzmodell. Die Archival Storage-Komponente wird durch den Tivoli Storage Manager (TSM) der Firma IBM realisiert und besteht aus einem Hierarchischen Speicher-Management-System (HMS), das neben Festplattenspeichern, die über das NFS-Protokoll mit dem LRZ verbunden sind, auch Tertiärspeicher in Form einer robotergesteuerten Medienbibliothek unterstützt. Datenströme, die auf ein Magnetband migriert werden, bleiben solange auf den Festplatten erhalten, bis neuer Speicherplatz benötigt wird. Dieses Caching steigert die Zugriffsperformance.

### 2.2.2 kopal

Das kopal-Projekt der Deutschen Nationalbibliothek (DNB), der Niedersächsischen Staats- und Universitätsbibliothek (SUB), der Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG) und der IBM wurde im Sommer 2007 abgeschlossen und hat gleichzeitig den Routinebetrieb aufgenommen. Es verfolgt ähnliche Ziele wie BABS und orientiert sich ebenfalls am OAIS-Referenzmodell.

Das IBM-Produkt DIAS (Digital Information Archiving System) bildet die Grundlage für ein OAIS-kompatibles Archiv. DIAS ist mandantenfähig, d.h. es kann von verschiedenen Organisationen gleichzeitig genutzt werden. Es unterstützt das *Universal Objekt Format* (UOF), welches auf den Formaten *Metadata Encoding and Transmission Standard* (METS) in der Version 1.4 und *Langzeitarchivierungs-Metadaten für Elektronische Ressourcen* (LMER) in der Version 1.2 basiert. Der Tivoli Storage Manager erlaubt die Verwendung verschiedener Speichermedien (Festplatten und Bänder) und die verteilte Speicherung von AIP. Außerdem ist es möglich, über standardisierte Schnittstellen und eigene Softwarekomponenten Archivpakete zu erstellen, einzuspielen und abzufragen. Hierzu wurde von der DNB und der SUB die Open-Source-Softwarebibliothek koLibRI (kopal Library for Retrieval and Ingest) entwickelt. DIAS wird im Rahmen von kopal ebenso wie der Archival Storage von der GWDG betrieben.

Im Gegensatz zu BABS verfügt DIAS über eine erweiterte Erhaltungsfunktionalität mit Hilfe des Konzeptes eines Universal Virtual Computers (UVC), das von IBM entwickelt wurde [IBM, 2008]. Durch eine Kombination von Emulation und Migration kann mit UVC ein hohes Maß an Technologieunabhängigkeit erreicht werden. In kopal wurde diese UVC-Funktionalität bisher nicht genutzt. Stattdessen wurden in koLibRI prototypische Komponenten zur Formatmigration implementiert, die auf Verwaltungsfunktionen von DIAS aufsetzen. Es ist hervorzuheben, dass es durch die klare Schnittstellenspezifikation von DIAS möglich ist, mit externen Komponenten auf die DIAS-Kernfunktion zuzugreifen und somit ein auf DIAS basierendes Archiv leicht in eine Grid-Umgebung zu integrieren.

### 2.2.3 Institutionelle Repositorien

Institutionelle Repositorien (IR) sind Archive, die Publikationen einer Universität oder Forschungseinrichtung sammeln, erhalten und zugänglich machen. Auch hier wird häufig das OAIS-Konzept zugrundegelegt. Im Gegensatz zu den beiden erstgenannten OAIS-Archiven steht hier aber die Langzeitarchivierung nicht im Vordergrund. Stattdessen sollen Forschungsergebnisse schnell und an zentraler Stelle entweder für Mitarbeiter oder auch für die Allgemeinheit (Open Access) publiziert werden. Zu den archivierten Materialien zählen – neben Open Access-Zeitschriften – Abschlussarbeiten, Technische Berichte oder Dissertati-

onen. Teilweise werden auch verwaltungstechnische Unterlagen wie Formulare, Kursverzeichnisse oder –beschreibungen über ein IR angeboten. Ein institutionelles Repository ist durch folgende Eigenschaften gekennzeichnet:

- Vorhalten von Inhalten (d.h. Dokumente und/oder Daten) an zentraler Stelle,
- globale Sichtbarkeit für eine (heute meist räumlich verteilte) Forschergemeinschaft oder für die Allgemeinheit,
- Indexierung der Inhalte mit Metadaten, deren Repräsentation in standardisierten Metadatenformaten erfolgt,
- Webschnittstelle für Ingest (Self-Archiving), Suchfunktionen sowie Access,
- Formatvalidierung und –migration sowie Integritätssicherung mittels Prüfsummen,
- SIPs und DIPs in standardisierten Containerformaten zur Unterstützung der OAIS-Funktionen durch externe Anwendungen.

Bekannte Beispiele für IRs sind Digitool, DSpace, Eprints, Fedora, und MyCoRe. Eine ausführliche, vergleichende Beschreibung dieser IRs findet man in [Borghoff et al., 2005]. Fedora bietet einen Webservice-basierten Zugriff auf die Archivfunktionen an. Es eignet sich daher auch besonders gut, um es ohne Webschnittstelle mit Hilfe von Computern anzusprechen.

Häufig erfüllen OAIS-Archive die Standards der Open Archive Initiative (OAI) und bieten u.a. auch die Möglichkeit, die Metadaten der gespeicherten Inhalte mit Hilfe des Protocol for Metadata Harvesting (OAI-PMH) zu exportieren.

Nach dieser Einführung in die LZA wollen wir uns dem Stand der Grid- und eScience-Technologien zuwenden.

### 3 Grid Computing und eScience

Während einem hohen Bedarf an Rechenleitung und Speicherkapazität in der Vergangenheit vorwiegend durch einzelne Supercomputer und Mainframe-Systeme begegnet wurde, werden heute häufig Verbundlösungen, bei denen mehrere Rechner über Datennetze miteinander kommunizieren, bevorzugt. Die Spannweite der eingesetzten Computer- und Netzwerkhardware reicht dabei von günstigen, in Massenfertigung hergestellten Komponenten bis hin zu hochleistungsfähigen Spezialanfertigungen. Eine standortübergreifende Kommunikation wird dabei im Allgemeinen nicht über dedizierte Festverbindungen, sondern über das Internet realisiert. Die Heterogenität der genutzten Ressourcen kann ebenso wie die Komplexität des Gesamtsystems durch eine als *Middleware* bezeichnete Softwareschicht vor dem Benutzer verborgen werden, der auf diesem Weg eine einheitliche Sicht auf alle ihm zur Verfügung stehenden Ressourcen erhält. In Abhängigkeit vom zugrunde liegenden Paradigma wird dabei zwischen Peer-to-Peer- (P2P) und Grid-Systemen unterschieden.

Peer-to-Peer-Systeme sind durch eine dezentrale Organisation gekennzeichnet, bei der die gemeinsam genutzten Ressourcen nicht auf wenige leistungsstarke Systeme konzentriert, sondern auf viele, im allgemeinen kostenfrei zur Verfügung gestellten Arbeitsplatzrechnern verteilt sind. Steinmetz und Wehrle [Steinmetz und Wehrle, 2004] definieren ein Peer-to-Peer-System als „ein sich selbst organisierendes System gleichberechtigter, autonomer Einheiten (Peers), das vorzugsweise ohne Nutzung zentraler Dienste auf der Basis eines Rechnernetzes mit dem Ziel der gegenseitigen Nutzung von Ressourcen operiert“. Durch den Verzicht auf

eine zentralisierte Infrastruktur können Flaschenhälse in Form von unterdimensionierten Netzwerkverbindungen oder überlasteten Server-Maschinen vermieden werden. Eine geographische Verteilung von gleichberechtigten Ressourcen verringert außerdem die Gefahr von Systemausfällen und Datenverlusten durch Naturkatastrophen, Sabotage-Aktionen, Unfälle oder lokale Konflikte. Des Weiteren entziehen sich die Inhalte von Staatsgrenzen überschreitenden P2P-Systemen der Kontrolle einzelner Regierungen, was unter Berücksichtigung entsprechender Sicherheitsvorkehrungen auch in Staaten ohne Recht auf freie Meinungsäußerungen eine ungehinderte Kommunikation ermöglicht.

Obwohl mit Hilfe des P2P-Paradigmas auf kostengünstigem Weg extrem leistungsstarke Systeme aufgebaut werden können, bringt die Nutzung einer großen Zahl von Endgeräten als Leistungsträger auch verschiedene technische und organisatorische Probleme mit sich, die vor einem Einsatz des P2P-Konzepts für das jeweilige Aufgabenfeld gelöst werden müssen. Die bereitstehenden Peers verfügen häufig über eine deutlich geringere Rechenleistung als geeignet dimensionierte Server-Systeme, was eine Verteilung von aufwändigen Berechnungen auf zahlreiche Peers erforderlich macht und einen hohen Verwaltungsaufwand nach sich zieht. Bei jeder Parallelisierung verbleibt jedoch ein sequentiell auszuführender Programmanteil, dessen Verarbeitungsdauer maßgeblich von der Leistungsfähigkeit der einzelnen Rechner abhängig ist. Die für die Netzwerkanbindung von Endgeräten üblichen niedrigen Übertragungsgeschwindigkeiten verlängern zudem die Antwortzeit für alle eintreffenden Benutzeranfragen. Ein weiteres zu lösendes Problem besteht in der Sicherung von Vertraulichkeit und Integrität, was u.a. durch geeignete Verschlüsselungsverfahren sowie den Einsatz von Prüfsummen erreicht wird. Weil Arbeitsplatzrechner im Gegensatz zu Server-Maschinen eher selten im Dauerbetrieb eingesetzt werden, ist die Verfügbarkeit der Ressourcen entsprechend gering, was insbesondere bei der Datenhaltung durch eine hinreichend große Zahl von Repliken kompensiert werden muss. Obwohl durch das Vorhalten entsprechend vieler Kopien die Gefahr von temporären Unverfügbarkeiten oder dauerhaften Datenverlusten erheblich verringert werden kann, ist eine permanente Erhaltung der Daten auf diesem Weg nicht zu gewährleisten. Trotzdem wird – wie nachfolgend anhand der Projekte LOCKSS und OceanStore exemplarisch angesprochen – auch das P2P-Paradigma gelegentlich für die Archivierung digitaler Daten verwendet.

LOCKSS steht für „Lots of Copies Keep Stuff Safe“ und bezeichnet eine gemeinnützige Initiative, um eine P2P-Software zu erstellen, die auf einem Netzwerk kostenlos bereitgestellter PCs (so genannte LOCKSS Boxen) digitale Inhalte langfristig speichert. Es können statische Dokumente (open access), Webseiten und Blogs gespeichert werden, die jeweils in die aktuell gültigen Formate migriert werden. Das Konzept von LOCKSS basiert auf der redundanten Speicherung von Repliken, deren Integrität durch fehlertolerierende Codes überprüft beziehungsweise im Fehlerfall auch wiederhergestellt werden kann. Auf der LOCKSS Box läuft eine Middleware, die das Web nach Inhalten durchsucht. Sofern sie von einem Herausgeber (überwiegend Universitäten) freigegebene Inhalte findet, werden diese in das verteilte Archiv aufgenommen. Die bestehenden Inhalte werden um Prüfsummen bzw. fehlertolerierende Codes ergänzt und durch eine Vielzahl von Kopien sorgt man für die nötige Redundanz zur langfristigen Archivierung. Der Zugriff erfolgt für den Benutzer transparent. Wenn ein Dokument eines Herausgebers nicht mehr direkt von dessen Server bereitgestellt werden kann, erhält der Benutzer das angeforderte Dokument über einen LOCKSS-Server, der als Proxy arbeitet. LOCKSS erlaubt lediglich die Archivierung freigegebener oder über das Web zugänglicher Inhalte. Alle Anbieter müssen ihre Inhalte bei LOCKSS anmelden bzw. registrieren. Benutzer können keine eigenen Inhalte einstellen und keine gezielte Suche (außer mit

einer Standard-Suchmaschine) durchführen. Außerdem bietet LOCKSS kein Rechtemanagement.

*OceanStore* ist ein auf dem P2P-Paradigma basierendes globales Dateisystem, das an der University of California entwickelt und mit Hilfe eines Prototyps, an dem sich über 40 Einrichtungen in Nordamerika, Europa und Australien beteiligt haben, getestet wurde. Die zu speichernden Daten werden gemeinsam mit verschiedenen für die Verarbeitung erforderlichen Informationen zu einem Datenobjekt verschmolzen, welches fragmentiert wird und über mehrere, ggf. sogar hunderte Rechner verteilt gespeichert wird. Die einzelnen Fragmente werden über eine baumartige Struktur verwaltet, in deren Wurzelknoten auch Metadaten abgelegt werden können. Aufgrund der verwendeten Codierung reicht bereits eine kleine Teilmenge der Fragmente aus, um den Inhalt einer Datei wieder vollständig herstellen zu können. Die verteilte Speicherung der Fragmente führt gemeinsam mit der fehlertoleranten Codierung zu einer weitgehenden Unempfindlichkeit gegenüber regionalen Krisen wie Naturkatastrophen oder militärische Konflikte. Die Hardware-Ressourcen eines OceanStore-Netzwerks werden in zwei Kategorien, so genannte Ringe, eingeteilt. Der innere Ring besteht aus einer Reihe leistungsstarker Server-Maschinen, die für die Speicherung der aktuellen Versionen aller Datenobjekte zuständig sind. Die Systeme des inneren Rings sollten vertrauenswürdig sein, da sie nicht nur für die Speicherung der aktuellen Datenobjekte, sondern auch für die Kontrolle der Schreibrechte verantwortlich sind. Sobald ein Client eine angeforderte Datei nicht in seinem lokalen Cache vorliegen hat, leitet er die Anfrage an die Maschinen des inneren Rings weiter. Sofern dort keine serialisierte Version der angeforderten Datei vorrätig ist, werden die Fragmente der Datei beschafft und serialisiert. Neben einer schnellen Internetanbindung benötigen die Maschinen des inneren Rings somit auch eine hohe Rechenleistung für die Serialisierung der Datenobjekte. Obwohl somit der Kern jedes OceanStore-Netzwerks durch leistungsstarke Server gebildet wird, unterliegt das Gesamtsystem keiner zentralen Kontrolle. Während die Maschinen des inneren Rings somit einer ganzen Reihe von Anforderungen genügen müssen, bestehen an die – im Allgemeinen von einzelnen Personen bereitgestellten – Rechner des äußeren Rings keine Mindestanforderungen. Diese Rechner werden für die Archivierung älterer Versionen der Datenobjekte verwendet, auf die nicht mehr schreibend zugegriffen werden kann.

Grid Computing wurde in den 90er Jahren des 20. Jahrhunderts durch die Idee motiviert, eine dem Stromnetz (grid) ähnlich Infrastruktur aufzubauen, die jedoch anstelle von elektrischer Energie typische EDV-Ressourcen wie z.B. Rechenleistung und Speicherkapazität bereitstellt. Gemäß dieser Grundidee sollte ein Benutzer seine Arbeitsaufträge dem Grid übergeben, welches dann eigenständig geeignete Ressourcen auswählt, die gesamte Verarbeitung steuert und überwacht und schließlich die erzielten Ergebnisse dem Benutzer zurückgibt und die beanspruchte Leistung abrechnet. Seitdem wurden zahlreiche Versuche zur Festlegung einer einheitlichen Definition des „grid“-Begriffs unternommen, was in der Grid Community zwar zu einem weitgehenden Konsens hinsichtlich der wesentlichen Merkmale von Grid-Systemen, aber zu keiner einheitlich akzeptierten Definition geführt hat (siehe hierzu auch Stockinger [Stockinger, 2007]). Eine kurze und prägnante Definition findet sich bei Stockinger [Stockinger, 2007] gemäß der ein Grid als „eine vollständig verteilte, dynamisch rekonfigurierbare, skalierbare und autonome Infrastruktur zur Bereitstellung eines ortsunabhängigen, überall vorhandenen, sicheren und effizienten Zugriff auf eine aufeinander abgestimmte Menge von Diensten zum Kapseln und Virtualisieren von Ressourcen (Rechenleistung, Speicherkapazität, Apparaturen, Daten etc.)...“ anzusehen ist. Insbesondere abstrahieren Grid-Systeme somit von technischen Details und ermöglichen eine gemeinsame Nutzung der eingebundenen Ressourcen, beispielsweise Rechenleistung und Speicherkapazität, durch eine große Anzahl

von Nutzer. In der Literatur werden Grid-Systeme häufig anhand der von ihnen hauptsächlich zur Verfügung gestellten Ressourcen in Rechen- (Compute Grid) und Daten-Grids (Data Grid) klassifiziert. Eine weitere wichtige Ressource des Grids stellt das Internet dar, das als globales Verbindungsnetzwerk dient. Aufgrund hohen Datenvolumens oder Engpässen der Übertragungsbandbreite kann die Performanz eines Grids sehr stark eingeschränkt werden.

Bei beiden Grid-Formen wird die angestrebte Abstraktion durch die Middleware hergestellt. Unter einer Middleware versteht man dabei eine zwischen Betriebssystem und Anwendungsprogrammen anzutreffende Software, die mit Hilfe einer Reihe von anwendungsneutralen Diensten und Protokollen die Kernfunktionen des jeweiligen Grid-Systems realisiert. Rechen- und Daten-Grids sind nach einem fünfstufigen Schichtenmodell aufgebaut. Auf der untersten Ebene sind die verwalteten Hardware-Komponenten, beispielsweise zur Datenerfassung, Speicherung oder Berechnung, angesiedelt, die dann zusammen mit der Systemsoftware die so genannte *Basic Grid Fabric* bilden. In einer darauf aufbauenden *Communication*-Schicht werden grundlegende Dienste zum Datentransfer (FTP bzw. GridFTP) zwischen zwei oder mehreren Knoten aus der untersten Schicht angeboten. Des Weiteren werden in dieser Schicht auch Dienste zum Datenschutz und zur Authentifikation der Benutzer realisiert. In der nächsten Ebene, der *Resource*-Schicht, werden sämtliche Dienste erbracht, die für eine gemeinsame Nutzung einzelner Ressourcen benötigt werden. Hierzu zählen unter anderem Managementfunktionen, welche die Konditionen für die Nutzung der Ressourcen (Zeitraum, QoS und Preis) aushandeln und die eigentliche Nutzung überwachen und steuern. Auf der nächst höheren, *Collective*-Schicht genannten Ebene werden sowohl das globale Ressourcenmanagement als auch die Wechselbeziehung zwischen mehreren Ressourcen koordiniert. *Resource*-Schicht und *Collective*-Schicht können zu einer einzigen Schicht zusammengefasst werden, die beispielsweise bei Venugopal et al. [Venugopal et al., 2006] als *Data Grid Services* bezeichnet wird und zur Verwaltung und zum transparenten Zugriff (z.B. OGSA-DAI) auf verteilt gespeicherte Daten dient. In der obersten, so genannten *Anwendungsebene* (application) werden schließlich die von den unteren Ebenen gelieferten Dienste den Benutzern zur Durchführung ihrer Aufträge bereitgestellt. Im Idealfall sind alle Zugriffe auf dieser Ebene bereits derart virtualisiert, dass sie der Benutzer wie bei der Nutzung von lokal verfügbaren Ressourcen verwendet kann.

Die Nutzergemeinschaften, die sich die Ressourcen teilen, verfolgen in der Regel ein gemeinsames Ziel, z.B. die Erforschung des Weltraums. Sie werden konzeptuell in so genannten Virtuellen Organisationen (VO) zusammengefasst und authentifizieren sich im Grid in der Regel mittels Zertifikaten. VOs können auch hierarchisch organisiert sein. Entsprechende Strukturen findet man dann auch in den gemeinsam genutzten Repositorien.

## **3.1 Grid Computing**

### **3.1.1 Rechen-Grids**

Unter einem Rechen-Grid (*Compute Grid*) versteht man „eine Hard- und Software-Infrastruktur, die einen zuverlässigen, konsistenten, permanenten und kostengünstigen Zugriff auf High End-Computersysteme“ [Foster und Kesselman, 1998] ermöglicht. Obwohl bei der Einrichtung eines Rechen-Grids die bereitgestellte Rechenleistung im Mittelpunkt des Interesses steht, wird ein hinreichend dimensioniertes Umfeld aus Netzwerkverbindungen, Speicherkapazitäten sowie Steuer- und Überwachungssoftware benötigt, um ein funktionsfähiges, eine große Zahl einzelner Rechner umfassendes Gesamtsystem zu verwirklichen. Die auf die-

se Weise zusammengeführten Ressourcen müssen alle eingehenden Arbeitsaufträge zügig und mit reproduzierbarem und nachhaltigem Ergebnis bearbeiten. Der Zugriff durch Benutzer und Administratoren erfolgt ebenso wie die Kommunikation zwischen den Diensten über standardisierte Schnittstellen und Protokolle, welche die vorhandene Heterogenität verbergen, ohne jedoch die verfügbare Rechenleistung herabzusetzen. Innerhalb des Grids sollten die dem Benutzer angebotenen Dienste dauerhaft verfügbar sein. Schließlich muss die Nutzung eines solchen Rechen-Grids auch vom ökonomischen Standpunkt gesehen attraktiv sein, weil die potenziellen Benutzer andernfalls die Anschaffung von eigenen Ressourcen präferieren.

In den letzten Jahren wurde eine Reihe von Middleware-Plattformen für Rechen-Grids entwickelt, welche die in der obigen Definition genannten Eigenschaften in mehr oder weniger großem Umfang auf unterschiedliche Weise umsetzen. Die Interoperabilität zwischen verschiedenen Grid-Middlewares sowie die Portabilität und Wiederverwendbarkeit einzelner Programme und Softwaresysteme kann durch die Einhaltung von Standards oder die explizite Spezifikation geeigneter Schnittstellen realisiert werden. Eine bedeutende Rolle spielt in diesem Zusammenhang die *Open Grid Services Architecture (OGSA)*, eine an Diensten orientierte Architektur, die neben einer Menge von Basisfunktionen auch deren Verhalten beschreibt.

Ein wesentliches Problem bei der Nutzung von Web-Diensten in Grids besteht darin, dass Web-Dienste zustandslos sind und somit keine Daten zwischen aufeinander folgenden Aufrufen des Dienstes ausgetauscht werden. Ein solcher Datenaustausch ist jedoch essentielle Voraussetzung für die Realisierung von aus mehreren Einzelaktionen bestehenden Funktionen. Durch die Nutzung von Datenbanken, temporären Zwischenspeichern oder geeigneten System- bzw. Grid-Diensten können Web-Dienste mit einem Zustand versehen werden. Das *Web Service Resource Framework (WSRF)* legt die benötigten Regeln für die Verwaltung und Nutzung der Zustandsinformationen und das Zusammenwirken mit zustandsbehafteten Ressourcen fest.

Gängige Middleware-Plattformen, wie das weit verbreitete *Globus Toolkit* (kurz: *Globus*) der Globus Alliance oder das am Forschungszentrum Jülich entwickelte *UNICORE*-System verwenden sowohl OGSA als auch WSRF zur Verbesserung der Interoperabilität mit anderen Grid-Technologien. In den nachfolgenden Unterabschnitten werden Globus, UNICORE und das vom Projekt *Enabling Grids for E-sciencE (EGEE)* entwickelte *gLite* stellvertretend für alle anderen Compute Grid-Plattformen behandelt.

### **3.1.1.1 Globus**

Die Entwicklung der Middleware *Globus* begann 1995 im Rahmen des Gemeinschaftsprojekts „Globus Project“, das 2003 in „Globus Alliance“ umbenannt wurde. Zu den ursprünglichen Projektpartnern, U.S. Argonne National Laboratory, Information Science Institute (ISI), University of Southern California und University of Chicago, kamen im Laufe der Zeit weitere Partner hinzu, darunter mit dem Royal Institute of Technology in Schweden und der University of Edinburgh zwei europäische Einrichtungen. Die Zielsetzung bei der Entwicklung bestand von Anfang an in der Bereitstellung einer grundlegenden Basisinfrastruktur sowie der darauf aufsetzenden Dienste zur Realisierung eines Rechen-Grids.

Globus ermöglicht die gemeinsame Nutzung verschiedener Ressourcen wie z.B. Rechenleistung und Datenbanken. Obwohl sich im Allgemeinen weder organisatorische noch geographische Grenzen auf diesen Vorgang auswirken, behalten alle teilnehmenden Einrichtungen letztlich die Hoheit über ihre eigenen Ressourcen. Die Besonderheiten der jeweiligen Hard- und

Softwareumgebungen werden von Globus vor den Benutzern verborgen. Um diese Aufgaben erfüllen zu können, stellt Globus zahlreiche Dienste und Bibliotheken bereit, deren Aufgaben von der Erkennung, Überwachung und Verwaltung der Ressourcen über die Erkennung von Fehlern bis hin zum Daten- und Kommunikationsmanagement reichen. Diese Dienste und Bibliotheken werden in Form einer Reihe von Software-Komponenten bereitgestellt, die sowohl gemeinsam als auch einzeln genutzt werden können. Eine Übersicht findet sich beispielsweise bei Foster [Foster, 2006].

Seit der Veröffentlichung der ersten Globus-Version im Jahr 1998 wurde die Middleware kontinuierlich weiterentwickelt und stellt in der gegenwärtig aktuellen Version 4.0 eine auf Web-Diensten basierende Plattform dar, die gegenüber den vorangegangenen Versionen eine deutlich höhere Stabilität, Geschwindigkeit und Bedienbarkeit sowie eine bessere Dokumentation, Standardkonformität und Funktionalität aufweist [Foster, 2006].

Globus dient als Grundlage für zahlreiche Community Grids auf der ganzen Welt. Die öffentliche Verfügbarkeit der Quelltexte und die Möglichkeit zur flexiblen Kombination der verschiedenen Komponenten von Globus ermöglicht eine einfache Adaption der Middleware an den Bedarf der jeweiligen Benutzergruppen. Durch die Einhaltung von Standards wie OGSA und WSRF besteht auch die Möglichkeit zur Interaktion mit Komponenten anderer Middleware-Plattformen. Entsprechend können auch Globus-Komponenten in Grid-Systemen, die auf einer anderen Middleware basieren, genutzt werden.

Nachfolgend werden mehrere Komponenten von Globus ausführlicher angesprochen, die für die im weiteren Verlauf dieser Expertise aufgezeigten Nutzungsmöglichkeiten von Grids im dLZA-Kontext Bedeutung haben. Dabei handelt es sich im einzelnen um das zwischenzeitlich in gsiFTP umbenannten GridFTP, die für die Integration von Datenbanken zuständige Schnittstelle Open Grid Service Architecture - Database Access and Integration services (OGSA-DAI) sowie um den Data Replication Service (DRS), der sich aus dem Reliable File Transfer Service (RFT) und dem Replica Location Service (RLS) zusammensetzt.

*GsiFTP (GridFTP)* ist eine auf den expliziten Transfer von Dateien in Grid-Systemen angepasste Version des in RFC 959 spezifizierten FTP-Protokolls. Es stellt einen sicheren und zuverlässigen Dienst für die Übertragung von Dateien zwischen Grid-Knoten zur Verfügung. GsiFTP erweitert das bekannte FTP-Protokoll um die Authentifizierung mit GSI (Grid Security Infrastructure), die ein Bestandteil des Globus Toolkit darstellt. GSI verwendet Zertifikate zur asymmetrischen Verschlüsselung der Datenübertragung und unterstützt Single Sign-On durch die Delegation-Methode mit einem Proxy. Der Proxy authentifiziert sich dazu stellvertretend für die Benutzer über ein neues, von diesem signiertes Zertifikat.

Das GsiFTP-Protokoll ermöglicht die Datenübertragung zwischen unterschiedlichen Rechnerplattformen. Um die Übertragungszeit zu verkürzen, können im Gegensatz zum normalen FTP gleichzeitig mehrere parallele Datenströme genutzt werden. Außerdem ist auch eine von dritter Seite („third party“) veranlasste Übertragung möglich, bei der ein Dateitransfer zwischen zwei Grid-Knoten von einem dritten Knoten angestoßen werden kann. Dies kann beispielsweise nötig sein, wenn die Eingabedaten für eine umfangreiche Berechnung an einem anderen Ort erzeugt wurden und zuvor erst auf den eigentlichen Rechenknoten übertragen werden müssen. Die Benutzer können diesen Transfer von einem beliebigen Standort aus beauftragen.

Die Abkürzung *OGSA-DAI* steht für *Open Grid Service Architecture - Database Access and Integration services*. Diese Middleware-Komponente ermöglicht den Zugriff auf Datenbanken und gehört ab Version 3.2 zum Umfang des Globus Toolkit. OGSA-DAI bietet Benutzern Grid-Dienste, um über WSRF-kompatible Web-Services transparent auf relationale und XML-basierte Datenbanken zuzugreifen. Von OGSA-DAI werden viele populäre Datenbanksysteme unterstützt. Neben Diensten zur Anfrage, Transformation und Bereitstellung von Daten wird auch ein Toolkit angeboten, um Client-Anwendungen zu erstellen.

Die Rechteverwaltung erfolgt auf der Benutzer-Ebene. OGSA-DAI setzt hierzu Benutzerzertifikate in Benutzerkennungen und Passworte für den Zugriff auf die Datenbanken um. Weitergehende Sicherheitsmechanismen müssen durch das Datenbanksystem implementiert werden. Außerdem findet keine vollständige Virtualisierung der Speicherressourcen statt, da die Benutzer geeignete Anfragen an die Datenbank schicken müssen, d.h. sie müssen deren Anfragesprache kennen, um auf die Daten zuzugreifen. Da OGSA-DAI modular implementiert ist, kann sein Funktionsumfang jedoch von Anwendern erweitert werden. In der neuesten Version OGSA-DAI 3.0 [Antonioletti, 2007] werden auch Workflows in Form verketteter Datenbankanfragen unterstützt. Hierdurch können zeitaufwendige Datentransfers vermieden werden.

Neben dem bereits angesprochenen GsiFTP verfügt Globus in der Version 4.0 über weitere Dienste zur Verwaltung und Übertragung von Daten im Grid. Obwohl das relativ einfache GsiFTP bereits für viele Aufgaben ausreicht, schränken verschiedene Eigenschaften des Protokolls dessen mögliche Einsatzgebiete ein. So erfordert der Einsatz von GridFTP beispielsweise einen offenen Socket auf Seiten des Clients, was zwar für einzelne kurze Übertragungen toleriert werden kann, bei länger dauernden oder sporadischen Datentransfers hingegen problematisch ist. Während Störungen im Netzwerk oder Ausfälle der Gegenseite mit Hilfe der im Hauptspeicher des Clients abgelegten Informationen keinen Neustart der unterbrochenen Übertragung erfordern, können die durch eine Funktionsstörung auf Seiten des Clients unterbrochenen Übertragungen nach dessen Neustart nicht wieder aufgesetzt werden, weil die hierfür erforderlichen Informationen nicht mehr verfügbar sind. Zur Lösung dieses Problems stellt Globus einen Web-Dienst namens *Reliable File Transfer (RFT)* bereit, der die Statusinformationen aller Übertragungen in einem permanenten Speicher vermerkt. RFT hinterlegt alle eintreffenden Anforderungen zur Durchführung von Datentransfers in einer Datenbank und startet die eigentlichen Übertragungen zu gegebener Zeit automatisch. Der Benutzer kann den aktuellen Stand seiner Übertragungen entweder über die Benutzerschnittstelle des Dienstes abfragen oder sich automatisch über andere WSRF-kompatible Globus-Dienste über auftretende Zustandsänderungen des Datentransfers informieren lassen.

Eine weitere in diesem Zusammenhang interessante Globus-Komponente ist der *Replica Location Service (RLS)*, mit dessen Hilfe die im Grid gespeicherten Dateien sowie deren Kopien verwaltet werden. Der RLS-Dienst basiert auf einem unter Umständen über mehrere Server und Standorte verteilten Register, in dem die physischen Speicherorte aller Dateien und ihrer Replikat verzeichnet sind. Neu erzeugte Dateien werden entweder manuell vom Benutzer oder automatisch durch einen anderen Grid-Dienst eingetragen. Mit Hilfe der auf diesem Weg hinterlegten Informationen können Grid-Benutzer zu einem späteren Zeitpunkt die vorhandenen Replikat einer Datei lokalisieren und verwenden.

Der *Data Replication Service (DRS)* ist ein zur WSRF-Spezifikation kompatibler Web-Dienst, der die beiden zuvor beschriebenen Dienste RFT und RLS verwendet, um für eine Gruppe von Dateien die Verfügbarkeit von jeweils einer Kopie an einem bestimmten Speiche-

rort des Grids zu gewährleisten. Für die Erledigung dieser Aufgabe wird zunächst mit Hilfe von RLS eine Kopie der gewünschten Datei ermittelt, die anschließend mit RFT an den betreffenden Zielort übertragen wird. Neu angelegte Kopien werden von DRS unter erneuter Verwendung des RLS-Dienstes als Replikat registriert.

### 3.1.1.2 UNICORE

*UNICORE* ist eine betriebsfertige Grid Middleware, deren Entwicklung 1997 am Forschungszentrum Jülich mit der Zielsetzung begann, den Nutzern der deutschen Supercomputer-Zentren einen sicheren und nahtlosen Zugriff auf die deutschlandweit verteilten heterogenen Ressourcen zu ermöglichen. In der gegenwärtig aktuellen Version UNICORE 6 wird dem Benutzer eine graphische Benutzerschnittstelle als Zugangspunkt zur Verfügung gestellt, über die sämtliche Funktionen von UNICORE in Anspruch genommen werden können. Technische und organisatorische Details der zugrunde liegenden Grid-Infrastruktur bleiben dem Benutzer dadurch verborgen.

Zum Funktionsumfang von UNICORE zählen nicht nur Dienste für die Realisierung eines sicheren und nahtlosen Zugriffs auf Daten, Hard- und Software-Ressourcen sondern auch für die Verwaltung der zu verarbeitenden Workflows und die Anbindung einer großen Spannweite von Ressourcen. UNICORE basiert auf OGSA und berücksichtigt zahlreiche Standards wie z.B. WSRF, mit deren Hilfe eine Zusammenarbeit mit anderen Middleware-Plattformen bzw. deren Komponenten erreicht wird. Daher kann an dieser Stelle auf eine detaillierte Behandlung der UNICORE-Architektur und einzelner Dienste verzichtet werden.

Zahlreiche nationale und internationale Projekte verwenden UNICORE als Middleware oder beteiligen sich an ihrer Weiterentwicklung bzw. Erweiterung. Beispiele hierfür sind AWARE, Chemomentum, DEISA, EUROGRID, VIOLA sowie Projekte der D-Grid-Initiative. Ebenso wie Globus wird UNICORE unter einer Open Source-Lizenz vertrieben, was eine Anpassung der Middleware an den individuellen Bedarf einzelner Communities ermöglicht.

### 3.1.1.3 gLite

Die Middleware *gLite* ist eine im Mai 2006 in der Version 3.0 freigegebene Entwicklung des Projekts *Enabling Grids for E-science (EGEE)*, die auf ihrem Vorgänger *European DataGrid (EDG)* und der *LHC Computing Project (LCG)* Middleware basiert. Das EGEE-Projekt vereint Wissenschaftler und Ingenieure aus über 40 Ländern, die gemeinsam an der Entwicklung einer nahtlosen Grid-Infrastruktur für das eScience-Umfeld forschen und arbeiten.

*gLite* setzt sich aus einem die Grundfunktionen realisierenden Middleware-Kern und eine Reihe höherer Dienste zusammen, die in Komponenten gebündelt sind und teilweise aus anderen Grid-Plattformen, beispielsweise Globus, übernommen wurden. Grid-Anwendungen erhalten dabei sowohl auf die höheren Dienste als auch die Basisfunktionen Zugriff. Gegenüber seinen Vorgängern EDG und LCG verfügt *gLite* unter anderem über verbesserte Schnittstellen zur Verwaltung von Datenbeständen und Jobs. Benutzer können alternativ über eine Kommandozeile oder ein Web-Portal auf das Grid zugreifen.

Die unter einer Open Source-Lizenz vertriebene gLite-Middleware wird häufig in Projekten der Hochenergiephysik eingesetzt, wobei eine zunehmende Verbreitung bei anderen Benutzergruppen mit einem hohen Bedarf an Rechenleistung, beispielsweise in der Astrophysik, zu beobachten ist.

### 3.1.2 Daten-Grids

Eine ausführliche und aktuelle Übersicht über Daten-Grids findet man in [Venugopal et al., 2006]. Dort wird auch eine Taxonomie angegeben. Daten-Grids dienen dazu, verteilte Datensammlungen (Repositorien) aus verschiedenen so genannten *Anwendungsdomänen* zusammenzuführen, um sie innerhalb von Nutzergemeinschaften (user communities) gemeinsam zu nutzen (sharing). Hierbei brauchen sich die Benutzer einer Domäne nicht bei den anderen Domänen anzumelden. Ein wesentliches Merkmal von Daten-Grids besteht darin, dass sie einen logischen Namensraum verwenden, um eine Vielzahl räumlich verteilter physischer Ressourcen in einem virtualisierten Dateisystem zu integrieren.

Die Funktionalität der weiter oben beschriebenen Schichtenstruktur wird bei einem Daten-Grid in einer entsprechenden Middleware gekapselt. Über diese kann ein Rechner um ein Betriebssystem- und ein Domänen-unabhängiges (Grid-)Dateisystem erweitert werden. Sobald die Benutzer sich mit dem Daten-Grid verbunden haben, können Teilverzeichnisse aus dem logischen Namensraum in das Dateisystem des eigenen Rechners transparent eingebunden werden. Die Middleware stellt andererseits aber auch dem Daten-Grid physische Speicherressourcen zur Verfügung, auf die dann von anderen Rechnern zugegriffen werden kann. Neben direktem Zugriff auf Festplattendateien wird auch die Einbindung von Datenbanken unterstützt. Zur Umsetzung logischer Dateinamen (bzw. -pfade) auf das physische Speichermedium in einer Ressource muss eine zentrale Verwaltung (registry) vorhanden sein, die quasi als Makler (broker) zwischen einer Anfrage und dem Anbieter fungiert.

In den nachfolgenden Unterabschnitten folgt eine kurze Vorstellung von mehreren Middleware-Systemen, die in der Praxis eine größere Bedeutung erlangt haben. Trotz der genannten Middleware-Systeme sind Daten-Grids heute immer noch unterrepräsentiert, obwohl sie sowohl für Rechen-Grids als auch eScience-Anwendungen von außerordentlicher Bedeutung sind.

#### 3.1.2.1 dCache

Das in Gemeinschaftsarbeit mehrerer europäischer und internationaler Forschungseinrichtungen<sup>1</sup> entwickelte dCACHE/SRM-System ist eine Middleware für Massenspeicher, die auf einer großen Palette unterschiedlicher Hardware- und Betriebssystemumgebungen lauffähig ist. Bei der Entwicklung von dCache wurde besonderer Wert auf eine gute Skalierbarkeit gelegt, die eine Nutzung sowohl in kleinen wie auch in sehr großen Speichersystemen ermöglicht. Die von dCache zur Verfügung gestellten Dienste und Schnittstellen wurden bewusst nicht auf den Bedarf einer einzigen Zielgruppe zugeschnitten, damit eine möglichst große Zahl von Fachgebieten von einem Einsatz der dCache-Middleware profitieren kann.

---

<sup>1</sup> Unter den beteiligten Einrichtungen befinden sich unter anderem das Deutsche Elektronen-Synchrotron (DESY) sowie das CERN.

Zum Funktionsumfang zählen neben der Verwaltung des verfügbaren Speicherplatzes eine verteilte, gegebenenfalls standortübergreifende Speicherung der Daten sowie die für den Benutzer transparente Durchführung der hierdurch erforderlich werdenden Datenübertragungen, des Weiteren die Unterstützung einer hierarchischen Speicherverwaltung (Hierarchical Storage Management) unter Berücksichtigung von Bandlaufwerken und den zugehörigen Roboter-Systemen. Die Steuerung von dCache erfolgt mit Hilfe einer Schnittstelle, die als Storage Resource Manager (SRM) bezeichnet wird.

Die Namen der gespeicherten Dateien werden bei dCache durch die Nutzung von Metadaten vom jeweiligen Speicherort getrennt. Durch das Anlegen von Replikaten kann zudem ein hohes Maß an Fehlertoleranz erreicht werden. Gemeinsam bewirken diese Eigenschaften von dCache, dass neue Knoten ohne Probleme nachträglich in ein bestehendes Daten-Grid integriert werden können und dass das Abschalten bzw. der Ausfall eines einzelnen Knotens für die Erhaltung der Daten unkritisch ist. In diesem Zusammenhang ist auch der Wegfall mehrerer Knoten unproblematisch, wenn vorher hinreichend viele Kopien angelegt wurden. Die Zahl der im Einzelfall anzulegenden Replikate kann durch die Festlegung einer Ober- und Untergrenze an den jeweiligen Bedarf angepasst werden.

Obwohl dCache die zu speichernden Daten automatisch den verfügbaren Ressourcen zuweist, kann die Verteilung der Daten durch entsprechende Vorgaben beeinflusst werden. Load-Balancing-Mechanismen gewährleisten eine Verteilung der Speicherlast gemäß einer konfigurierbaren Strategie. Die verteilte Speicherung von Daten wird durch verschiedene Maßnahmen, wie zum Beispiel dem verzögerten Zugriff auf transiente Daten, vor den Benutzern verborgen. Dateizugriffe erfolgen entweder über eine Art URI (Universal Resource Identifier) oder mit Hilfe verschiedener Derivate von FTP-Diensten.

Sicherheit wird bei dCache durch den Einsatz des Authentifizierungsdienstes Kerberos sowie durch die Nutzung verschiedener SSL-Protokolle angestrebt. Durch die Integration weiterer Module kann dCache um zusätzliche Schutzmaßnahmen erweitert werden

dCache stellt nur geringe Anforderungen an die Hardware der Speichersysteme, weshalb es mit einer großen Zahl von unterschiedlichen Hardwareplattformen und Speichermedien zusammenarbeiten kann. Die Verwendung von Standard-Protokollen macht dCache von bestimmten Einsatzgebieten unabhängig und erleichtert die Interaktion mit anderen Speichersystemen. dCache ist unter anderem kompatibel zu gsiFTP (GridFTP), SRM und LDAP und verfügt über Schnittstellen zu gängigen Speichermanagern wie zum Beispiel dem Tivoli Storage Manager (TSM), Open Storage Manager (OSM) und High Performance Storage System (HPSS).

dCache wird inzwischen in Systemen eingesetzt, deren Datenaufkommen zum Teil bei mehreren hundert Terabyte pro Tag liegt.

### **3.1.2.2 SRB**

Storage Resource Broker (SRB) ist eine autonom lauffähige Middleware für Daten-Grids, die seit 1995 am San Diego Supercomputer Center (SDSC) entwickelt wird. SRB verbindet heterogene Datenbestände unter einer einheitlichen Benutzerschnittstelle, wobei es im kombinierten Einsatz mit einem Metadatenkatalog (MCAT) den ortsunabhängigen Zugriff auf Spei-

cherressourcen, Daten und Metadaten gewährleistet. Durch die Nutzung von Metadaten, die unter anderem Angaben zu Dateien, Benutzern, Benutzergruppen und Ressourcen umfassen können, besteht die Möglichkeit zur Realisierung komplexer Suchfunktionen, die dem Nutzer den Zugriff auf die im Grid verteilten Daten erleichtern. Daten können somit allein anhand ihrer Attribute gefunden werden.

Zum Funktionsumfang des SRBs zählen neben den bereits genannten Punkten die Bereitstellung von Diensten zur Replikation, Datensicherung und Datenwiederherstellung. Des Weiteren verfügt SRB auch über Basisdienste zur Erfassung, Verwaltung und Archivierung der Daten.

Durch die Verwendung eines einheitlichen globalen Namensraums können die gespeicherten Objekte eindeutig identifiziert werden, was die Zusammenarbeit an gemeinsam genutzten Daten deutlich erleichtert. Die gespeicherten Daten können durch die Verwendung so genannter Collections und Subcollections hierarchisch organisiert werden. SRB optimiert den Zugriff auf transient gehaltene Daten, indem es die entstehenden Latenzen durch Datenkompression, Zwischenspeicher (Cache-Speicher), Container zur Zusammenfassung mehrerer kleiner Datenobjekte und den Zugriff auf geeignet platzierte Replikate reduziert. Durch den Einsatz dieser Maßnahmen kann SRB Dateizugriffe durchführen, deren Geschwindigkeit bei großen Dateien höher als bei Nutzung von FTP ist. Beim Zugriff auf kleinere Dateien treten hingegen wegen der Nutzung von MCAT geringe Verzögerungen auf.

Sicherheit wird bei SRB unter anderem durch Verfahren zur Authentifikation, Autorisierung, Verschlüsselung und Zugriffskontrolle angestrebt. Dabei greift SRB sowohl auf die Schutzmaßnahmen des für den MCAT-Dienst verantwortlichen DBMS zurück, als auch die Sicherheitsvorkehrungen, die auf den für Datenhaltung zuständigen Systemen getroffen wurden. Die Authentifikation wird alternativ über die Client-Software oder unter Verwendung der Grid Security Infrastructure (GSI) von Globus realisiert. Zum Schutz des Netzwerkverkehrs wird Encrypt1, eine schnelles, Passwort-basierendes Verfahren zur Autorisierung, verwendet.

SRB unterstützt die meisten unter Windows und Unix verfügbaren Dateisysteme, Speicher-Managementsysteme wie z.B. HPSS, verbreitete Datenbanken wie DB2 und Oracle sowie Band-Archivierungssysteme. Als Betriebssystem wird auf den einzelnen Grid-Knoten Windows, Mac OS-X oder eines der zahlreichen UNIX-Derivate (AIX, Linux, Sun Solaris etc.) vorausgesetzt.

Die SRB-Middleware wird bei einer großen Zahl von Einrichtungen und Projekten, beispielsweise dem National Virtual Observatory (NVO), dem Biomedical Informatics Research Network (BIRN) oder dem Hayden Planetarium Visualization Project, eingesetzt.

### **3.1.2.3 Nirvana**

Nirvana ist eine kommerzielle Variante des im vorigen Unterabschnitt behandelten SRB, deren Weiterentwicklung durch die Nirvana Division von General Atomics erfolgt. Während die vom San Diego Supercomputer Center gepflegte SRB-Version primär auf die Bedürfnisse und Anforderungen eines wissenschaftlichen Umfelds ausgerichtet ist, zielt die von Nirvana vertriebene Variante hauptsächlich auf Kunden in Industrie und Kommerz ab, die auf die Verfügbarkeit einer Service-Infrastruktur angewiesen sind.

Die Koexistenz zweier unterschiedlicher SRB-Versionen erklärt sich somit durch die Notwendigkeit, verschiedene Benutzerkreise mit unterschiedlichen Bedürfnissen bedienen zu müssen. Obwohl beide Entwicklerteams sowohl hinsichtlich der Planung als auch der Umsetzung neuer Funktionen weiterhin zusammenarbeiten, ist aufgrund der unterschiedlichen Zielgruppen ein kontinuierliches Auseinanderdriften beider Varianten zu beobachten.

#### **3.1.2.4 iRODS**

iRODS (i Rule Oriented Data Systems) ist eine adaptive Grid-Middleware, die seit Herbst 2004 am San Diego Supercomputer Center (SDSC) entwickelt wird. Im Gegensatz zu SRB, bei dem die Regeln für das Management der Daten hart im System codiert werden und Anpassungen an den individuellen Bedarf der Benutzer nur durch die Konfiguration von Policies möglich ist, können bei iRODS durch die Programmierung von Regeln eine feingranulare Anpassung des Systems an den individuellen Bedarf der jeweiligen Benutzer vorgenommen werden. Diese Anpassung erfolgt mit Hilfe des so genannten Rule Oriented Programmings (ROP), bei dem durch geeignete Regeln festgelegt wird, unter welchen Bedingungen bestimmte Operationen, sog. Micro-Services, ausgeführt werden. Sobald die Voraussetzungen für eine Regel vorliegen, wird diese aktiv und die zugehörigen Mikro-Services werden ausgeführt. Die Verarbeitung der Daten kann dabei durch eine Umgestaltung bzw. den Austausch der Regeln oder eine Veränderung der darunter liegenden Micro-Services beeinflusst werden.

Ein auf iRODS basierendes Daten-Grid setzt sich aus heterogenen verteilten Speicher- und Rechnersystemen zusammen, deren Tätigkeit durch eine Software zur Auswertung und Verarbeitung der Regelmengen geleitet wird. Die Verwaltung der den Daten anhängenden Attribute erfolgt ebenso wie die Überwachung des Zustands der Daten sowie der auf ihnen ausgeführten Operationen mit Hilfe einer Datenbank.

Im Gegensatz zu den in den beiden vorangegangenen Unterabschnitten behandelten Middleware-Systemen ist die Zahl der von iRODS unterstützten Plattformen, Betriebssystemen und Protokollen zurzeit noch sehr gering. Obwohl das dargestellte Konzept von iRODS insbesondere hinsichtlich der Adaptivität an neue Systemumgebungen und Aufgabenstellung interessante Konzepte für eine zukünftige Daten-Grid Middleware aufweist, kann die tatsächliche Leistungsfähigkeit dieses Systems zum gegenwärtigen Zeitpunkt noch nicht beurteilt werden, weil sich iRODS noch in der Alpha-Phase befindet und wenige im praktischen Einsatz befindlichen Referenzinstallationen existieren.

Tabelle 1 gibt einen Überblick über die in diesem Abschnitt behandelten SRB-Varianten unter Berücksichtigung der öffentlich zugänglichen Informationen zur Rechtelage und den entstehenden Kosten.

Tabelle 1: Übersicht über die verfügbaren SRB-Versionen

SRB-Version	Hersteller	Homepage	Rechtelage und Kosten
SDSC-SRB	SDSC	<a href="http://www.sdsc.edu/srb/">http://www.sdsc.edu/srb/</a>	kostenlos für Lehre und Forschung, Konditionen für kommerzielle Nutzung auf Anfrage
Nirvana-SRB	General Atomics	<a href="http://www.nirvanastorage.com/">http://www.nirvanastorage.com/</a>	Kommerzielle Software, Konditionen auf Anfrage
iRODS (0.9.2)	SDSC	<a href="http://www.irods.org">http://www.irods.org</a>	BSD-Lizenz, kostenlos, Open Source-Software

### 3.2 eScience

Zu eScience zählt man sowohl die orts- und/oder zeitunabhängige Zusammenarbeit in Kerngebieten der Wissenschaft als auch die dazu benötigten Werkzeuge und Infrastrukturen. Eine geeignete Infrastruktur für eScience-Projekte besteht aus einer verteilten, im Allgemeinen globalen Systemumgebung, mit deren Hilfe neben dem Austausch von Daten und Nachrichten auch die gemeinsame Nutzung der angebotenen Instrumente und Rechnersysteme möglich ist. Für die Realisierung einer solchen Infrastruktur wird neben geeigneten Rechner- und Netzwerkressourcen eine Software benötigt, die eine (geographisch) verteilte Steuerung und Verwaltung sämtlicher Ressourcen realisiert, den Austausch von Daten ermöglicht sowie ggf. Funktionen für die Abrechnung der in Anspruch genommenen Leistungen bereitstellt. Des Weiteren sollte durch geeignete Schutzmaßnahmen ein für den jeweiligen Zweck hinreichendes Maß an Vertraulichkeit, Integrität, Authentizität und Verfügbarkeit gewährleistet werden.

Mit Hilfe der Grid-Technologie werden bereits seit mehreren Jahren erfolgreich weltumspannende Rechnerverbünde mit der Zielsetzung aufgebaut, den teilnehmenden Benutzern eine gemeinsame Nutzung der angebotenen Ressourcen zu ermöglichen, wodurch nicht nur die anfallenden Arbeitsaufträge schneller erledigt, sondern auch die angebotenen Ressourcen mit einer höheren Effizienz genutzt werden. Aufgrund der Verwendung des Internets als Weitverkehrsmedium müssen Grid-Systeme über geeignete Mechanismen zum Schutz von Arbeitsaufträgen, Daten und Benutzer verfügen. Verfahren für die Realisierung eines Abrechnungswesens sind ebenfalls verfügbar.

Die offensichtliche Korrelation zwischen den für die eScience formulierten Anforderungen und den für die Grid-Technologie angeführten Eigenschaften legt bereits eine Nutzung von Grid-Systemen zur Realisierung einer eScience Infrastruktur nahe. In Deutschland wird seit dem 1. September 2005 im Rahmen der D-Grid-Initiative eine Grundlage für den Einsatz von Grid-basierenden eScience-Technologien in unterschiedlichen Wissensgebieten, wie z.B. der Astronomie, Hochenergiephysik oder Medizin, geschaffen. An eine gemeinsam genutzte generische Grid-Middleware werden die Middlewares der partizipierenden Community Grids angebunden und auf diesem Weg eine disziplinübergreifende Zusammenarbeit ermöglicht. Als Plattformen werden Globus, UNICORE und gLite unterstützt. Für die Archivierung von Projektergebnissen und wichtigen Datenbeständen existiert derzeit keine einheitlich von allen D-Grid-Projekten genutzte Plattform.

Auch im internationalen Umfeld findet man viele erfolgreiche Verbände von Wissenschaftlern verschiedener (meist naturwissenschaftlicher) Disziplinen, die auf Basis von Grids zusammenarbeiten und darüber hinaus im Laufe der Zeit kollaborative Umgebungen zur gemeinsamen Nutzung von teuren Spezialgeräten (wie z.B. Teleskopen) oder zum Daten- und Informationsaustausch entwickelt haben. Hier findet man Synergiepotenziale, die insbesondere für experimentell arbeitende Forschergruppen von großer Bedeutung sind. Durch die Fernsteuerung von Experimenten wird es nicht nur möglich, Kosten einzusparen, sondern auch die teuren Spezialgeräte besser auszulasten. Ein Beispiel für solche auch als Collaboratives [Bos et al., 2007] bezeichneten eScience-Technologien ist die International Virtual Observatory Alliance (IVOA), die im Bereich Astronomie Standards für Datenübertragungsprotokolle und astronomische Datenarchive entwickelt hat [Quinn, 2004].

Neben den im D-Grid zusammengefassten eScience-Projekten, die eine Grid-Infrastruktur für die Umsetzung ihrer Forschungsvorhaben verwenden, gibt es eine Reihe von D-Grid-Partnerprojekten aus den Bereichen eLearning und Wissensmanagement, deren Fokus in der Entwicklung von Informations-, Kommunikations- oder Publikationsplattformen besteht bzw. die für bestimmte Bereiche eine langfristige Verfügbarkeit wissenschaftlicher Daten gewährleisten sollen. Diese Projekte können einerseits einen ergänzenden Beitrag für die Einrichtung einer einheitlichen und allgemein verwendbaren Kommunikationsstruktur anderer eScience-Projekte leisten und andererseits von der Rechenleistung und den Speicherkapazitäten einer Grid-Infrastruktur sowie der Verfügbarkeit eines ebenfalls angebotenen OAIS-kompatiblen Langzeitarchivs profitieren. In den nachfolgenden Unterabschnitten werden exemplarisch die Projekte eSciDoc, ONTOVERSE, WIKINGER und Im Wissensnetz beschrieben, da diese zahlreiche kollaborative und wissenserzeugende Komponenten enthalten, die aus der eScience-Perspektive mögliche Synergiepotenziale bieten.

### **3.2.1 eSciDoc**

Das vom Bundesministerium für Bildung und Forschung (BMBF) geförderte und von den Projektpartnern Max-Planck-Institut und Fachinformationszentrum Karlsruhe vorangetriebene eSciDoc-Projekt dient der Entwicklung einer integrierten Informations-, Kommunikations- und Publikationsplattform für netzbasiertes wissenschaftliches Arbeiten, mit dessen Hilfe die wissenschaftliche Kooperation im eScience-Umfeld unterstützt werden soll. Als weiteres Ergebnis dieses bis Juli 2009 laufenden Projekts wird die langfristige Verfügbarkeit von Forschungsergebnissen und –materialien der Max-Planck-Gesellschaft angestrebt.

Die Umsetzung des Projekts erfolgt – soweit wie möglich – durch Nutzung vorhandener Open Source-Softwarelösungen unter Berücksichtigung internationaler Konzepte und Standards. Durch die Bereitstellung von umfangreichen Schnittstellen und APIs soll die Interoperabilität mit einer großen Zahl anderer Systeme gewährleistet werden.

eSciDoc setzt sich aus einem Framework und einer Reihe von Anwendungen zusammen. Die Aufgabe des Frameworks besteht in der Realisierung von Diensten, welche für die Abbildung beliebiger Dokumente sowie der dazu gehörenden Kontext-Informationen, über den gesamten Lebenszyklus der Dokumente hinweg, benötigt werden. Die angebotenen Dienste stellen hierfür Funktionen zur Verwaltung von Metadaten und Formaten und zum In- und Export von Daten sowie verschiedene Suchfunktionen zur Verfügung. Das Framework setzt sich aus mehreren zentralen Komponenten – darunter ein Repository, eine Prozesssteuerung

und eine Rechteverwaltung – zusammen, die bei der Erledigung der anfallenden Aufgaben zusammenwirken. eSciDoc umfasst eine so genannte Scholarly Workbench zur Unterstützung der Arbeit mit digitalen Datenbeständen, ein für die Verwaltung, Verbreitung und Archivierung wissenschaftlicher Arbeiten zuständiges Publication Management, eine eLib genannte Anwendung zur Sammlung und Archivierung elektronischer Verlagspublikationen und Datenbanken sowie ein elektronisches Labor-Journal namens eLab, das der dauerhaften Sicherung der aus Experimenten gewonnenen Ergebnisse dient.

Das Gesamtsystem ist aufgrund seines modularen Aufbaus und der Nutzung etablierter Konzepte und Standards hochgradig flexibel hinsichtlich zukünftiger Erweiterungen und Anpassungen. Ein Einsatz in anderen Anwendungsgebieten und eine damit verbundene Nachnutzung der entwickelten Komponenten wird seitens des Projekts ebenso wie eine Zusammenarbeit mit anderen Projekten – insbesondere aus den Themengebieten Langzeitarchivierung und Grid-Systeme – angestrebt.

### **3.2.2 ONTOVERSE**

Das vom BMBF geförderte ONTOVERSE-Projekt<sup>2</sup> befasst sich mit der Entwicklung einer Umgebung zur über Disziplingrenzen hinweg greifenden Erstellung von Ontologien, an deren Aufbau sich Forscher aus der ganzen Welt beteiligen können. Auf diesem Weg soll die gegenwärtige Vorgehensweise, bei der Ontologien dezentral von einzelnen Wissenschaftlern entworfen werden und häufig nur ein eingeschränkter Austausch mit anderen Forschern erfolgt, optimiert und beschleunigt werden.

Die Benutzerschnittstelle von ONTOVERSE wird in Form eines Wikis realisiert, das die interaktive Kooperation unter den beteiligten Wissenschaftlern erleichtern soll. Dabei soll die Kommunikation zwischen den beteiligten Personen durch den Einsatz eines Annotationsystems unterstützt werden. Zu den von ONTOVERSE angebotenen Leistungen zählen Hilfsmittel zur Vereinigung vormals eigenständiger Ontologien, Tools zur Visualisierung sowie Zugriffs- und Navigationsfunktionen. Weil der Erfolg des ONTOVERSE-Systems maßgeblich von einer regen Beteiligung der betroffenen Wissenschaftler abhängig ist, werden außerdem Dienste bereitgestellt, die eine Bildung von Arbeitsgruppen unterstützen.

Im Rahmen von ONTOVERSE sollen Grid-Technologien zur Erlangung einer zuverlässigen und effizienten Nutzung von Rechen-, Speicher- und Übertragungsleistungen sowie zur Sicherstellung einer hohen Skalierbarkeit und Verfügbarkeit eingesetzt werden. Durch Virtualisierung werden bei ONTOVERSE verteilte heterogene Ressourcen vereinheitlicht. Die Laufzeit des Projekts endet im September 2008.

### **3.2.3 WIKINGER**

Ziel des vom BMBF geförderten Projekts WIKINGER ist die Schaffung einer integrierten Plattform, deren Dienste eine effiziente und ortsunabhängige Recherche in Wissensbasen interessierender Fachgebiete ermöglichen, die Generierung neuen Wissens durch Zusammenar-

---

<sup>2</sup> ONTOVERSE ist ein aus den Begriffen **Ontologie** und **Universe** zusammengesetztes Kunstwort.

beit von Wissenschaftlern über das Internet unterstützen und eine Vernetzung neuer Beiträge mit bereits vorhandenem Wissen erleichtern.

Um diese Ziele zu erreichen, kombiniert WIKINGER verschiedene bekannte Technologien, die es um neue Methoden ergänzt. Die Vernetzung der gespeicherten Daten erfolgt halbautomatisch, indem zunächst Personen-, Funktions- und Ortsnamen aus vorhandenen Datenbeständen extrahiert und anschließend durch einen Computer miteinander verknüpft werden. Die dabei eingesetzte Software zur Erkennung von Eigennamen muss ein hohes Maß an Toleranz gegenüber unerwarteten Eingabedaten aufweisen und an neue Themengebiete adaptierbar sein. Rückmeldungen der Benutzer führen zu Anpassungen des erzeugten Netzes, welches auf diesem Weg auch an sich ändernde Interessen und Themenschwerpunkte angepasst werden kann. Die semantische Verknüpfung der gespeicherten Objekte ist eine wesentliche Voraussetzung für die Realisierung einer thematischen anstelle einer Stichwort-basierten Suchfunktion.

Die Benutzerschnittstelle von WIKINGER besteht aus einem Wiki-System, das um zusätzliche Funktionen für den Einsatz mit semantischen Netzen erweitert wurde. Die Verwendung von Wikis erlaubt eine dynamische Veränderung von Inhalten durch mehrere Autoren und ermöglicht dadurch die Entstehung von Wissen unter Mitarbeit aller beteiligten Forscher. Zudem fördern Wikis die Akzeptanz des Gesamtsystems aufgrund einer gegenüber anderen Konzepten kurzen Einarbeitungszeit sowie ihrer im Allgemeinen hohen Benutzerfreundlichkeit.

Die ersten Untersuchungen werden exemplarisch auf dem Themengebiet der „Katholischen Zeitgeschichte“ durchgeführt. Für den späteren Praxiseinsatz bietet sich die Nutzung von Grid-Infrastrukturen an, weshalb seitens des Projekts Aktivitäten in den Bereichen „Web-Services“ und „Open Grid Service Architecture“ (OGSA) geplant sind. Der Abschluss des Projekts ist für September 2008 geplant.

### **3.2.4 Im Wissensnetz**

Der Ausgangspunkt des vom BMBF geförderten Projekts „Im Wissensnetz“ liegt in der Beobachtung, dass übliche Ansätze zum Wissensmanagement in der Forschung keine Anwendung finden, weil Forschungsaktivitäten hochgradig dynamisch sind und noch kein allgemein akzeptiertes Vokabular etabliert werden konnte. Die unzureichend ausgeprägte Zusammenarbeit von Wissenschaftlern führt insbesondere in interdisziplinären Forschungsbereichen zu einem unvollständigen Kenntnisstand der einzelnen Forscher über bereits gewonnene Erkenntnisse der Kollegen und deren Kompetenzen.

Aufgrund dieser Umstände besteht das zentrale Ziel des Projekts in der Verbesserung der für die Schöpfung neuen Wissens notwendigen Prozesse durch den Einsatz von eScience-Werkzeugen. Dabei wird wissenschaftliches Arbeiten als eine Vernetzung von Lernprozessen verstanden, bei der die beteiligten Personen mit ihrem Wissen ebenso wie der gegenseitige Wissensaustausch im Mittelpunkt stehen. „Im Wissensnetz“ soll diesen Vorgang durch eine Verbesserung des Networkings zwischen Personen unterstützen und deren Zusammenwirken zur Lösung von wissenschaftlichen Problemen fördern. Durch die automatische Verknüpfung vorhandenen Wissens soll der daraus zu erzielende Nutzen gesteigert werden.

Um das gegebene Ziel erreichen zu können, werden unter anderem die Schaffung einer Sammlung von Werkzeugen zur Unterstützung der wissenschaftlichen Arbeit (Semantic Desktop), die Realisierung einer Möglichkeit zur Verknüpfung mehrerer Semantic Desktops innerhalb eine Gemeinschaft nach dem Peer-to-Peer-Ansatz und die Förderung von kooperativen Innovationsprozessen durch intelligente Assistenz-Programme als Ergebnisse angestrebt. Die Vernetzung einzelner Personen und Personengruppen soll dabei durch Dienste im Stile von Xing (ehemals openBC) sowie durch Dienste zur automatischen Erkennung möglicher Ansätze für Zusammenarbeit und zur halbautomatischen Erzeugung von Personenprofilen unterstützt werden. Inhalte werden durch die Analyse von Texten und die Nutzung von semantisch aufgewerteten Suchmaschinen vernetzt.

Der Abschluss des Projekts „Im Wissensnetz“ ist für Oktober 2008 geplant. Vorzeitig veröffentlichte Prototypen werden bereits von einzelnen Anwendern, vorwiegend Fraunhofer-Instituten, eingesetzt.

## **4 Befragung von drei CG-Projekten**

Im Rahmen der vorliegenden Studie wurde eine größere Zahl von nationalen, europäischen und internationalen Projekten aus den Bereichen eScience, Grid und Langzeitarchivierung hinsichtlich möglicher Synergiepotenziale untersucht und auf Basis der dabei gewonnen Erkenntnisse die in den nachfolgenden Kapitel dargestellten Synergiepotenziale und Handlungsempfehlungen erarbeitet. Außerdem wurden mit AstroGrid-D, C3-Grid und TextGrid drei Projekte der deutschen D-Grid-Initiative exemplarisch hinsichtlich der derzeit genutzten Archivierungsverfahren und des in dieser Hinsicht bestehenden bzw. für die Zukunft absehbaren Handlungsbedarfs befragt. In den nachfolgenden Abschnitten werden die genannten Projekte und die bei der Befragung erzielten Ergebnisse kurz dargestellt. Für die Zusammenarbeit möchten wir uns an dieser Stelle bei den befragten Projekt-Beteiligten bedanken.

### **4.1 AstroGrid-D**

Das Projekt AstroGrid-D soll eine landesweite, auf dem Globus-Toolkit 4 aufbauende Grid-basierende Infrastruktur schaffen, in die neben Rechner-Ressourcen auch astronomische Instrumente und Datenarchive eingebunden werden sollen. Mit Hilfe dieser Infrastruktur wird sowohl eine Verbesserung der Zusammenarbeit innerhalb der nationalen als auch der internationalen astronomischen Wissenschaftsgemeinde angestrebt. Dabei sollen die notwendigen Eigenentwicklungen des Projekts möglichst offen gehalten werden, damit auch andere Forschergruppen einen Nutzen aus dem investierten Entwicklungsaufwand ziehen können. Zu den vorgesehenen Einsatzgebieten zählen im Wesentlichen Simulationen größeren Ausmaßes sowie die Nachbearbeitung der dabei erzielten Ergebnisse, umfangreiche Analysenvorgänge auf zum Teil mehreren Terabyte großen verteilten Datenbeständen sowie die Anbindung spezieller astronomischer Instrumente [Enke et al., 2007].

Im Rahmen dieser Aufgaben werden häufig Datenbestände im Umfang von mehreren Terabyte verarbeitet, die aufgrund ihrer Größe im Allgemeinen weder mittel- noch langfristig von einzelnen HPC-Einrichtungen gespeichert oder sogar archiviert werden. Als Folge müssen die erzeugten Datenmengen an andere Speicherorte übertragen werden, was von AstroGrid-D zukünftig durch ein geeignetes Datei- und Metadatenmanagement unterstützt werden

soll. Simulationen, für deren Ausführung eine große Zahl von Prozessoren benötigt werden, profitieren außerdem aufgrund der höheren Zahl verfügbarer Rechner von der in Entwicklung befindlichen Grid-Infrastruktur. Eine Vernetzung der weltweit verteilten Teleskope ermöglicht nicht nur ein besseres Ressourcenmanagement mit einer effizienteren Nutzung, sondern auch eine deutlich höhere Chance für eine erfolgreiche Durchführung einer Beobachtung. AstroGrid-D plant die Entwicklung einer einheitlichen Schnittstelle für die Nutzung und Verwaltung (Überwachung und Scheduling) von robotischen Teleskopen.

Unsere Befragung des AstroGrid-D-Projekts hat ergeben, dass als Hilfsmittel für eine gemeinsame Datennutzung das Globus Toolkit Vers. 4 und Webservices, wie z.B. ein selbstentwickeltes Datenstrommanagement, eingesetzt werden. Ein im Rahmen von kollaborativen Aktivitäten erforderlicher Zugriff auf große Datensätze anderer Wissenschaftler erfolgt auf klassischem Weg, beispielsweise über das File Transfer Protokoll (FTP) bzw. unter Verwendung von Grid-Methoden wie z.B. GsiFTP oder OGSA-DAI. Die kollaborative Auswertung von Daten wird durch die geringe Zahl der gegenwärtig in Datenbanken gespeicherten Datensätze und das Fehlen eines einheitlichen Datensatzformates erschwert. Eine Daten- bzw. Dokumenten-Organisation zur langfristigen Archivierung von Datensätzen oder Dokumenten wird seitens des Projekts nicht betrieben.

Die gegenwärtig in der Astronomie anzutreffende Infrastruktur der im Bereich der Datenarchivierung aktiven Organisationen wird seitens des Projekts als nicht tragfähig und eine Ergänzung um eine Grid-Infrastruktur als sinnvoll erachtet. AstroGrid-D arbeitet in diesem Zusammenhang an der Entwicklung geeigneter Verfahren zur Nutzung von Grid-Technologien und der Lösung der aus der Heterogenität von Daten und Metadaten resultierenden Probleme. Die benötigte Speicherkapazität wird selbst nach wissenschaftlicher Auswertung der Daten noch einige Terabyte pro Tag betragen, wobei jedoch weder das Datenformat noch der genaue Umfang der zu archivierenden Daten bislang definiert wurden. Eine Einrichtung eigener Archive ist derzeit nicht geplant, weshalb die Langzeitarchivierung von bereits existierenden Archiv-Einrichtungen geleistet werden soll. Dementsprechend wurden seitens des Projekts keine konkreten Vorgehensweisen für die Ein- bzw. Auslieferung der zu archivierenden Daten angegeben. Die Suche nach einzelnen Datensätzen bzw. Dokumenten soll über die Grenzen einzelner Archive hinweg anhand von einzelnen Objektinformationen, wie z.B. den Koordinaten, physikalische Eigenschaften und referenzierende Publikationen, sowie nach Replikaten möglich sein. Obwohl derzeit sowohl eine graphische Benutzeroberfläche als auch ein Zugriff per Kommandozeile angeboten werden, wird für die Realisierung einer Grid-Technologien verwendenden LZA-Funktion eine programmierbare Schnittstelle als sinnvoll angesehen. Hinsichtlich der von den Nutzern akzeptierten Zugriffszeit konnten keine allgemeingültigen Angaben gemacht werden. Von einer LZA-Unterstützung wird erwartet, dass sie zum integrierten Bestandteil einer eScience-Umgebung mit Service-orientierter Funktionalität wird. Die zu entwickelnden Softwaremodule sollen als OpenSource veröffentlicht werden. Als vordringlich zu realisierende LZA-Funktionalität wurde die Entwicklung eines Communities übergreifenden Konzepts zur Realisierung einer LZA und eine Beschreibung der hierfür erforderlichen Dienste benannt.

## **4.2 C3-Grid**

*Collaborative Climate Community Data and Processing Grid (C3-Grid)* ist ein Projekt im Rahmen der D-Grid-Initiative, das den im Bereich der Klimaforschung tätigen Wissenschaftlern eine Grid-basierte Arbeitsumgebung zur Verfügung stellen soll, mit deren Hilfe über ein

einheitliches Portal sowohl der Zugriff und die Verarbeitung der verteilt gespeicherten heterogenen Datenbestände der Community als auch deren Verwaltung auf einfache Weise ermöglicht wird [Ulbrich et al., 2007]. Für die Realisierung eines einheitlichen Zugriffs auf die Datenbestände verschiedener Anbieter bedarf es eines gemeinsamen Metadatenschemas, das von C3-Grid unter Berücksichtigung der ISO-Richtlinien 19115 und ISO 19139 im Jahr 2006 standardisiert und um C3-Grid-spezifische Anpassungen ergänzt wurde. Anbieter von Metadaten müssen ihre eigenen Spezifikationen diesem Standard anpassen, um in das C3-Grid integriert werden zu können. Die Umstellung der Metadaten kann entweder von Hand oder automatisch erfolgen.

Die Interaktion zwischen Benutzer und Grid erfolgt über ein auf dem GridSphere Framework aufbauendes Portal, welches über den *Data Information Service* (DIS) und den *Workflow Scheduling Service* (WSS) mit der übrigen Grid-Infrastruktur kommuniziert. Während der erstgenannte Dienst anhand von Metadaten die dazugehörigen Daten innerhalb des Grids ermittelt, ermöglicht der letztgenannte die Verarbeitung von Arbeitsaufträgen, die sich aus mehreren Einzelschritten zusammensetzen können. In beiden Fällen werden die angeforderten Daten über das *Data Management System* (DMS) von C3-Grid zur Verfügung gestellt, das mit Hilfe einer einheitlichen Schnittstelle die Eigenschaften der verwendeten Speicherressourcen vor dem übrigen Grid verbirgt. Für die Ausführung der vom WSS verplanten Workflows werden Rechner-Ressourcen mit lokalen Speichern (*Workspaces*) in die C3-Grid-Infrastruktur eingebunden. Die für den jeweiligen Arbeitsschritt benötigten Daten werden mittels GridFTP aus den Archiven in die lokalen Speicher übertragen, wobei im gesamten Grid ein einheitlicher logischer Namensraum erhalten bleibt. Eine ausführlichere Darstellung der C3-Grid-Architektur findet sich bei Hiller und Fritzsich [Hiller und Fritzsich, 2007].

Als Middleware wird ein auf den besonderen Bedarf des C3-Grids angepasstes Globus Toolkit Version 4 verwendet. Allgemein wird die Einhaltung gängiger Standards und die Nutzung der vom D-Grid unterstützten Software-Komponenten angestrebt und die Entwicklung proprietärer Lösungen so weit wie möglich umgangen.

Aufgrund des hohen Datenaufkommens können nicht alle in der Klimaforschung erzeugten Daten langfristig erhalten werden. Aus diesem Grund wurde beispielsweise am World Data Center for Climate (WDCC) und am Deutschen Klimarechenzentrum (DKRZ) ein Konzept für die Langzeitarchivierung der zu erhaltenden Daten entwickelt, bei dem die langfristig zu speichernden Daten von den nur kurzfristig benötigten (Projekt-)Datenbeständen separiert werden. Testdaten werden dabei nur für wenige Wochen oder Monate, Projektdaten für wenige Jahre und abschließende Ergebnisse für mindestens zehn Jahre aufbewahrt. Bei den längerfristig zu speichernden Daten wird ein besonderes Augenmerk auf die dazugehörige Dokumentation gelegt, da eine sinnvolle Nutzung dieser Daten andernfalls fraglich ist. Gemäß den spezifischen Anforderungen der drei genannten Kategorien werden die Daten in einer vierstufigen Speicherhierarchie abgelegt, bei der unter Berücksichtigung von Zugriffsgeschwindigkeit und Speicherkapazität geeignete Medien für das Abspeichern der Daten gewählt werden. Eine detaillierte Darstellung dieses Konzepts findet sich bei Lautenschlager und Stahl [Lautenschlager und Stahl, 2007]. Wie bereits in der Studie von Klump [Klump, 2007] beschrieben wurde, werden die veröffentlichten Daten mit persistenten Identifikatoren versehen, die einen gezielten Zugriff auf die in Publikationen verwendeten Daten ermöglichen.

Unsere Fragen hinsichtlich des derzeitigen Entwicklungsstandes und der Anforderungen an eine zukünftige LZA-Funktion wurden sowohl seitens des C3-Projekts als auch vom WDCC und dem World Data Center for Remote Sensing of the Atmosphere (WDC-RSAT) beantwor-

tet. Die nachfolgende Darstellung fasst die von den drei Projektpartnern gegebenen Antworten in verkürzter Form zusammen.

Die gemeinsame Nutzung von Daten erfolgt über das Webportal des C3-Grids. Die Anbieter der Daten stellen diese dem Grid in Form von ISO 19115 Templates zur Verfügung, wo sie gemeinsam mit den entsprechenden Daten der anderen Anbieter in einem zentralen Datenkatalog zusammengefasst werden und für Suchanfragen verfügbar sind. Der Zugriff auf die eigentlichen Daten erfolgt über eine standardisierte Schnittstelle mit Hilfe von GsiFTP. Der für eine Zusammenarbeit erforderliche Zugriff auf große Datensätze von Kollegen erfolgt für C3-Grid-Benutzer über das Portal von C3-Grid; innerhalb von WDCC und WDC-RSAT können direkte Zugriffe auf Datenbanken vorgenommen werden, wobei das WDCC eine Oracle-Datenbank und das WDC-RSAT eine Flat File-Datenbank verwendet. Für eine kollaborative Auswertung der Datensätze stellt C3-Grid zum jetzigen Zeitpunkt erst einige vorgefertigte Workflow-Prototypen zur Verfügung. Eine langfristige Speicherung sowie eine Archivierung ist zum gegenwärtigen Zeitpunkt noch nicht Gegenstand von C3-Grid, sondern liegt im Verantwortungsbereich der jeweiligen Datenanbieter. Der von WDCC im Laufe der letzten 15 Jahre gesammelte Datenbestand umfasst derzeit ca. 340 Terabyte, zu denen aktuell zwischen 60 und 100 Terabyte jährlich dazukommen. In Zukunft ist mit einer Steigerung von etwa einem halben bis einem Petabyte pro Jahr zu rechnen. Die Auswahl des zu archivierenden Materials geschieht anhand von Forderungen aus wissenschaftlichen Programmen und Projekten sowie anhand von wissenschaftlichen Entscheidungen der Benutzer. Die Erfassung der Metadaten erfolgt über XML-Templates nach ISO 19115. Institutionelle Repositorien wie dspace, eprints oder fedora werden nicht verwendet. Alle Daten liegen in digitaler Form vor. Die benötigten Speicherressourcen werden im Rahmen des C3-Grid-Projekts zur Verfügung gestellt, die durch institutionelle Speicher der kooperierenden ICSU WDCs und des DWD ergänzt werden. Eine Dokumentation der im Rahmen des CG realisierten LZA-Funktionen liegt derzeit noch in den Anfängen. Der für Aufbau und Betrieb einer LZA-Infrastruktur investierte Personalaufwand beträgt beim WDC-RSAT circa 0,4 Personenjahre, beim WDCC etwa 5 Personenjahre.

Die in der Zukunft benötigten Speicherkapazitäten werden voraussichtlich insgesamt im Bereich von einigen Petabyte liegen, wobei allein die Zahl der im WDCC gespeicherten Datenentitäten derzeit im Bereich von  $5,8 \cdot 10^9$  liegt. Für die Zukunft werden Zuwachsraten in etwa derselben Größenordnung erwartet. Die Ein- bzw. Auslieferung der Datensätze zur Langzeitarchivierung soll über gängige Standard-Protokolle wie FTP, GridFTP oder http erfolgen. Als Suchkriterien sollen u.a. die geographische Region, Zeitpunkt, Parameter der Messung (z.B. die Temperatur) sowie Angaben zum zugrunde liegenden Experiment herangezogen werden können. Während für das C3-Grid-Portal eine graphische Benutzerschnittstelle realisiert ist, wird für den Zugriff auf die Datenarchive eine Script-basierte Schnittstelle bevorzugt. Eine Zugriffszeit von mehreren Minuten wird als akzeptabel erachtet, wobei diese Dauer beim WDC-RSAT deutlich unterschritten wird. Datenarchive wie die WDCs dienen bereits heute als Gedächtnisorganisationen des C3-Grids. Für die Zukunft ist der Aufbau eines C3-Grid Repository Centers (C3RC) geplant, das auch für das Grid selber die Frage der Langzeitarchivierung relevant werden lässt. Als besondere Anforderung an eine LZA-Unterstützung wird dabei insbesondere eine Unterstützung bei der Definition des Ingest-Prozesses gewünscht.

### 4.3 TextGrid

Das im Februar 2006 gestartete Projekt *TextGrid* soll die Arbeit von Geisteswissenschaftlern an historischen deutschen Texten durch die Nutzung von Grid-Technologien unterstützen. Historische Texte werden zu diesem Zweck mit hoher Qualität eingescannt und das dabei entstehenden Datenaufkommen über TextGrid in einem Daten-Grid abgespeichert. Durch die Verbindung dieser digitalen Abbildungen mit für Computer lesbaren Textdokumenten werden neue Möglichkeiten zur Verarbeitung und Auswertung der gespeicherten Dokumente geschaffen. TextGrid unterstützt sowohl den Aufbau als auch die Erweiterung einer solchen Verknüpfungsstruktur durch die teilweise Automatisierung der hierfür erforderlichen Arbeitsschritte.

Die graphische Benutzerschnittstelle soll alle technischen und organisatorischen Details verbergen, die nicht für die eigentliche Anfrage von Bedeutung sind. Für die Nutzung der Daten werden unterschiedliche Sichten angeboten. Arbeitsabläufe, die sich aus mehreren Einzelschritten zusammensetzen, sollen durch die graphische Benutzerschnittstelle intuitiv gesteuert bzw. modifiziert werden können. Der Zugriff kann alternativ über ein Web-Portal oder eine dedizierte Zugangssoftware erfolgen.

Durch die Einhaltung internationaler Standards und die Nutzung von Open Source-Produkten soll ein offener, virtueller und Grid-basierter Arbeitsbereich geschaffen werden, in dem die Verarbeitung, Analyse, Verknüpfung und Publikation von Texten ermöglicht wird. Die interne Darstellung der Texte erfolgt mittels des in den Geisteswissenschaften weit verbreiteten Standards *Text Encoding Initiative* (TEI). TextGrid, welches auf dem Globus-Toolkit in der Version 4.0 aufsetzt, ist modular aufgebaut und kann aufgrund der bereits erwähnten Konformität zu gängigen Standards sowie durch die Offenlegung der projektspezifischen Module von projektfremden Entwicklern durch das Einpflegen weiterer Module um zusätzliche Funktionen ergänzt werden. Ebenso sollen Benutzer das System durch die Integration geeigneter Module an den eigenen Bedarf anpassen können.

TextGrid plant eine Einbindung von Archiven, die auf diesem Weg für die Grid-Umgebung nutzbar gemacht werden. Es wird eine offene Plattform angestrebt, an die sich jedes interessierte Text-Archiv anschließen und die von jedem Forschungsprojekt genutzt werden kann [Gietz et al., 2006].

Auf unsere Fragen nach den derzeit verwendeten Werkzeugen zur gemeinsamen Nutzung von Daten wurde auf das Globus Toolkit Version 4.0 und die im Rahmen des TextGrid-Projekts entwickelten Hilfsmittel verwiesen. Für die kollaborative Auswertung der Datensätze stellt TextGrid eine eigene Arbeitsumgebung (Virtual Research Environment) zur Verfügung. Die Langzeitarchivierung wird als eines der Kernziele von TextGrid gesehen, wobei jedoch noch keine Preservation Policy vorliegt. Das bis Projektende für die Langzeitarchivierung erwartete Datenvolumen liegt bei etwa 12 Terabyte – für die nächsten 5 Jahre wird mit einem Aufkommen von 40 - 50 Terabyte gerechnet. Die Speicherressourcen werden im Rahmen des D-Grid-Verbundes bereitgestellt und von TextGrid anteilig mit betrieben.

An eine zukünftige LZA-Funktion werden seitens der TextGrid-Gemeinschaft die folgenden Anforderungen gestellt: Die Ein- bzw. Auslieferung der Daten an die LZA soll über

TextGrid-Utilities oder einer Massendaten-Schnittstelle möglich sein. Die bereitzustellenden Suchfunktionen sollen sowohl eine Volltextsuche als auch eine gezielte Suche innerhalb der Metadaten unterstützen. Eine Verwendung der XML Query Language auf XML/TEI-Annotationen sollte möglich sein. Für die Nutzung der LZA-Funktionen wird eine graphische Benutzerschnittstelle gewünscht. Obwohl TextGrid auch selbst am Aufbau eigener LZA-Repositories interessiert ist, ist die Nutzung von LZA-Diensten bestehender Gedächtnisorganisationen geplant. Die neue LZA-Funktion sollte eine langfristige Archivierung gewährleisten können, modular und offen sein. Als vordringlich zu realisierende LZA-Ziele werden die Integration von Repositorien in Grid-Umgebungen, die Einführung und Nutzung persistenter Identifikatoren sowie die Bereitstellung von LZA-Diensten wie z.B. Dateiformat-Verzeichnisse, Format-Validierung oder die Beobachtung von technologischen Entwicklungen genannt.

#### **4.4 Zusammenfassung**

Wie die vorangegangenen Abschnitte gezeigt haben, weisen die drei befragten Projekte neben diversen Unterschieden auch eine ganze Reihe von Gemeinsamkeiten hinsichtlich der bisher vorliegenden Realisierung und der für die Zukunft geplanten LZA-Aktivitäten auf. Tabelle 2 gibt einen Überblick über die von den drei Projekten bei der Bestandsaufnahme gegebenen Antworten. Bei den mit „-/-“ markierten Zellen wurden seitens des Projekts keine Angaben gemacht.

Während Tabelle 2 Auskunft über den Status quo der betrachteten Projekte hinsichtlich der im Kontext von Kollaboration und Langzeitarchivierung relevanten Methoden und Verfahren gibt, erlauben die in Tabelle 3 zusammengefassten Antworten einen Überblick über die Anforderungen der Projekte an eine zukünftige LZA-Funktion. Bei sämtlichen mit „-/-“ markierten Zellen wurden seitens des Projekts keine Angaben gemacht.

Beide Tabellen geben die Antworten der Projekte zum Teil in deutlich verkürzter Form wieder. Auf unsere Bitte um Benennung von im Fragebogen nicht angesprochenen wichtigen Aspekten wurde von keinem der drei Projekte geantwortet.

Tabelle 2: Bestandsaufnahme über die derzeit im LZA-Kontext eingesetzten Werkzeuge und realisierten Strukturen

	<b>AstroGrid-D</b>	<b>C3-Grid</b>	<b>TextGrid</b>
Verwendete Middleware	Globus Toolkit 4.X	Globus Toolkit 4.X	Globus Toolkit 4.X
Verfügbare Werkzeuge zur/für den ...			
... gemeinsamen Datennutzung	Globus-Werkzeuge und Webservices	Portal des C3-Grids	TextGrid- und Globus-Werkzeuge
... Zugriff auf Daten von Kollegen	FTP, GsiFTP, Webservices, OGSA-DAI, ...	Portal des C3-Grids; intern über DB-Zugriffe	TextGrid-Werkzeuge
... gemeinsamen Auswertung von Datensätzen	in Entwicklung	Vorgefertigte Workflows (Prototypen / in Entwicklung)	TextGrid-Werkzeuge
Angaben zur Archivierung von Daten oder Dokumenten:	(nicht vorhanden)	(Aufgabe der Datenanbieter)	
- Zeitraum, in dem die Daten gesammelt wurden	-/-	15 Jahre (WDCC)	-/-
- Umfang des bisherigen Bestandes	-/-	Massenspeicher 5 PB, Datenbank 340 TB (WDCC)	6 TB
- Auswahl des zu archivierenden Materials	-/-	Forderungen aus wissenschaftlichen Programmen und Projekten	-/-
- Erfassung der Metadaten	-/-	XML-Templates nach ISO 19115	automatisch und manuell
- Nutzung institutioneller Repositorien	-/-	TIB-Order	geplant
- Form der Daten	-/-	digital	-/-
- Betreiber der Speicherressourcen	-/-	C3-Grid, ICSU WDCs und DWD	D-Grid
- Dokumentation der LZA-Funktionen	-/-	in Entwicklung	-/-
- bisheriger Personalaufwand zur Realisierung von LZA-Funktionen	-/-	0,4PJ (WDC-RSAT) 5 PJ (WDCC)	-/-

Tabelle 3: Anforderungen der Projekte an eine zukünftige LZA-Funktion

	<b>AstroGrid-D</b>	<b>C3-Grid</b>	<b>TextGrid</b>
Erwarteter Bedarf an Speicherkapazität	einige TB pro Tag	mehrere Petabyte	40 – 50 TB
Erwartete Zahl von Datensätzen / Entitäten	-/-	$5,8 \cdot 10^9$ pro Jahr	-/-
Hilfsmittel zur Ein-/Auslieferung der Datensätze und Dokumente bei der LZA	nicht festgelegt, da kein eigenes Archiv geplant	Standard-Protokolle wie FTP, GridFTP oder http	TextGrid-Werkzeuge / Massendaten-Schnittstelle
Vom Archiv bereitzustellende Suchfunktionen / zu berücksichtigende Suchkriterien	Objektinformationen, referenzierende Publikationen, Replikate; Suche soll über Archivgrenzen hinweg möglich sein	Informationen zum Experiment und zu Parametern der Messung	Volltextsuche, Durchsuchung der Metadaten, Anfragen auf XML/TEI-Annotationen
gewünschte Benutzerschnittstelle	graphisch und programmierbar	Portal: graphisch Datenbank: programmierbar	graphisch
akzeptierte Zugriffszeiten	keine allgemeingültigen Angaben möglich	mehrere Minuten	-/-
Nutzung der von Gedächtnisorganisationen angebotenen Dienstleistung	noch nicht festgelegt	Nutzung der WDCs; Aufbau eines C3-Grid Repository Centers (C3RC) geplant; gemeinsamer Katalog vorhanden	Nutzung und Bereitstellung von Dienstleistung ist geplant
Anforderungen an eine LZA-Unterstützung	integrierter Bestandteil einer eScience-Umgebung mit Service-orientierter Funktionalität; Open Source	Unterstützung bei der Definition des Ingest-Prozesses	langfristige Archivierung, modular und offen
vordringlich zu realisierende LZA-Funktionalitäten	Entwicklung eines Community-übergreifenden Konzepts; Beschreibung der erforderlichen Dienste	Unterstützung bei der Definition und Realisierung des Ingest-Prozesses in C3RC	Integration von Repositorien in Grid-Umgebungen; Einführung und Nutzung persistenter Identifikatoren und Bereitstellung von LZA-Diensten

## 5 Synergiepotenziale

Synergie bezeichnet das Zusammenwirken im Sinne von „sich gegenseitig fördern“. Durch dieses Zusammenwirken entsteht ein Ganzes, das mehr als die Summe seiner Teile ist. Im Rahmen dieser Expertise sollten diese Synergiepotenziale zwischen Grid-/eScience-Technologien und der LZA identifiziert werden.

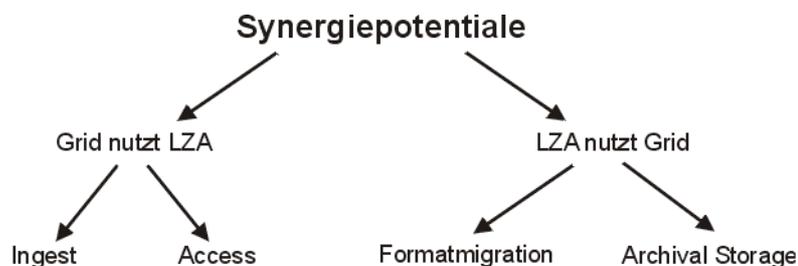
Wie aus der Bestandsaufnahme der CGs und Abschnitt 3.2 hervorgeht, stellen Grid-Technologien die Infrastruktur für eScience zur Verfügung. Neben dem Zugriff auf Rechnerressourcen ist auch der Austausch von Dokumenten und Daten für die Wissenschaftler von großer Bedeutung. Die Daten können entweder bei Experimenten oder Simulationen anfallen und sind meist sehr umfangreich [Klump, 2007]. Daher ist es teilweise unumgänglich, diese Daten im Grid verteilt zu speichern. Wie die CGs angeben, wird zum Datentransfer zwischen den Benutzerrechnern bzw. Grid-Knoten auf standardisierte Software wie FTP, http und GridFTP/GsiFTP zurückgegriffen. Wie die Analyse der befragten CGs weiter zeigt, haben diese fachspezifische eScience-Infrastrukturen aufgebaut, die sowohl aus Grid-Ressourcen als auch aus z.T. umfangreichen Archiven bestehen. Dabei können im Wesentlichen Dokument- und Datenarchive unterschieden werden.

In TextGrid werden digitale Dokumente und Digitalisate verarbeitet, die zusammen mit den Metadaten in proprietären Repositorien gesammelt werden. Die Archivobjekte und Metadaten liegen in standardisierten Formaten vor, die zur LZA in OAIS-kompatible Archive von Gedächtnisorganisationen eingelagert werden sollen. Zur mittelfristigen Archivierung werden auch institutionelle Repositorien in Betracht gezogen. Nach Auskunft der CG werden die eingelagerten Inhalte aus den Langzeitarchiven der Gedächtnisorganisationen nur abgefragt, wenn Arbeitskopien in den lokalen Repositorien unbrauchbar werden. Statt der originalen Digitalisate verwendet man als Arbeitskopien oft Formate mit einem geringerm Speicherbedarf (und damit auch geringerer Qualität) wie z. B. JPEG. Für die zu archivierenden Dokumente werden dagegen die Dokumentformate mit höchster Qualität (z. B. TIFF) eingesetzt.

Bei AstroGrid und C3-Grid spielen Daten von Simulationen und Beobachtungen bzw. Messungen eine wichtige Rolle (vgl. auch [Klump, 2007]). Im Laufe der letzten Jahre entstanden umfangreiche Archive in der Größenordnung von einigen Hundert Terabytes. Wie in Lautenschlager und Stahl [Lautenschlager und Stahl, 2007] eindrucksvoll gezeigt wird, steigen wegen der immer leistungstärkeren Computersysteme die benötigten Speicherkapazitäten dramatisch an (beim WDCC hat allein die Datenbank einen Umfang von ca. 340 TB). Dies führt schließlich dazu, dass die Kosten für die Speicherressourcen die Kosten für Rechnerressourcen übersteigen. Um dieser Entwicklung entgegenzuwirken, wurden Speicherbereiche mit unterschiedlichen Erhaltungshorizonten eingeführt. Nur in der höchsten Kategorie (docu) erfolgt beim WDCC eine LZA für Zeiträume von mehr als 10 Jahren. Die realisierten Datenarchive bestehen im Wesentlichen aus einer Datenbank und hierarchisch organisierten Speichersystemen, die neben Festplatten vor allem Magnetbänder zur LZA verwenden. Beim WDCC müssen neben den zu archivierenden Daten auch beschreibende Metadaten nach den ISO 19115- und 19139-Standard durch die Benutzer bereitgestellt werden. Die Daten selbst werden im standardisierten netCDF-Format gespeichert (<http://www.unidata.ucar.edu/software/netcdf/docs/>). Für die Bearbeitung codierter Daten gibt es frei zugängliche Tools, die bei Änderung der Codierung migriert werden und stets zu alten Codierungen abwärtskompatibel sind.

Im Hinblick auf die vorhandenen umfangreichen Bestände ist es verständlich, dass die CGs sich eine nahtlose Integration der existierenden Archive wünschen. Außerdem sollen sowohl die existierenden als auch neu hinzukommenden Archive über eine einheitliche Schnittstelle aus dem Grid heraus zugreifbar sein. Hierdurch würde für die CGs ein deutlicher Mehrwert aufgrund der Bereitstellung von LZA-Funktionen entstehen. Wie weiter oben festgestellt wurde, bedeutet Synergie aber auch, dass im Gegenzug die Archive von Grid-Diensten profitieren sollen.

In Abbildung 2 sind die Synergiepotenziale entsprechend den gegenseitigen Nutzungsmöglichkeiten in einer Übersicht dargestellt. Bei den Bezeichnungen gehen wir von Archivprozessen aus, die sich am OAIS-Referenzmodell orientieren. Im Idealfall ergibt sich bei jedem Archivprozess ein Mehrwert für das Grid bzw. die Archive. Aus der Sicht des Grids sind Mehrwerte für die Synergiepotenziale Ingest und Access erreichbar. Die Archive können vor allem von rechenintensiveren Formatvalidierungen bzw. -migrationen durch Rechner- und Speicherressourcen des Grids profitieren.



**Abbildung 2: Synergiepotenziale**

Obwohl es denkbar wäre, sämtliche technischen Archivprozesse im Grid zu implementieren, raten wir von einer solchen vollständigen „Gridifizierung“ ab. Durch eine verteilte Implementierung auf Grid-Ressourcen ohne einheitliche Administration können Sicherheitsrisiken und Instabilitäten beim Betrieb entstehen. Es ist vielmehr zu empfehlen, dass Archive durch Gedächtnisorganisationen oder Daten-Provider betrieben werden, die durch enge Kopplung der technischen wie organisatorischen Archivprozesse eine bessere Verfügbarkeit und Erhaltungssicherheit gewährleisten können. Insofern stünde allerdings einer Gridifizierung eines Archivs durch ausschließlich eigengenutzte Grid-Ressourcen nichts entgegen.

Die zu integrierenden Archive stellen wie die Grid-Ressourcen in sich geschlossene, aber im allgemeinen heterogene Systeme dar, die über das Internet lose miteinander gekoppelt sind. Um solche Systeme untereinander zu verbinden, bieten sich serviceorientierte Architekturen (SOA) an. Obwohl bei einer SOA die unterschiedlichsten Protokolle für die Abstimmung zwischen den Anbietern und Nutzern eines Dienstes herangezogen werden können, haben sich im Grid-Bereich die auf international anerkannten Standards beruhenden Webservices [Foster, 2005] zum de facto-Standard für den Informationsaustausch zwischen Diensteanbietern und -nutzern entwickelt. Diese Webservices sollten daher auch für die Anbindung von Archiven an eine Grid-Infrastruktur verwendet werden.

Ein Webservice erlaubt die Interoperabilität heterogener Systeme, da er eine plattform- und sprachunabhängige Nutzung von verteilten Anwendungen ermöglicht. Motiviert durch die Vorarbeiten der befragten CGs und des WDCG könnte die in Abbildung 3 skizzierte Architektur als Grundlage für eine Grid-basierte LZA-Infrastruktur dienen.

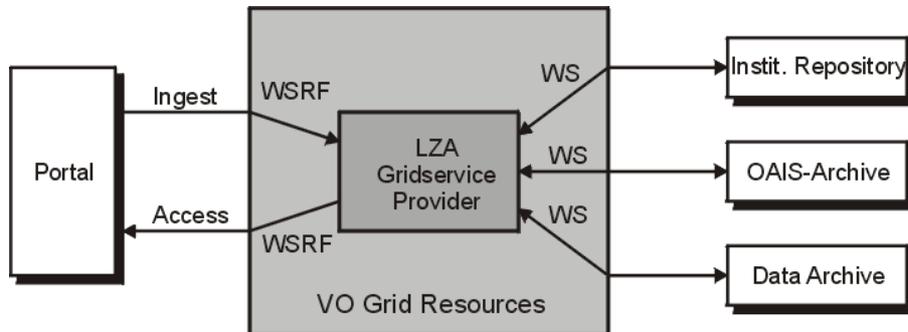


Abbildung 3: Serviceorientierte LZA-Infrastruktur (WS=Webservice)

Den Kern bildet ein LZA-Grid-Service-Provider, der auf einer Grid-Ressource der VO läuft und Webservices zur LZA anbietet. Über einen Verzeichnisdienst (Universal Description, Discovery and Integration protocol, UDDI) kann der Provider die von ihm bereitgestellten Dienste veröffentlichen und die Dienstenutzer können damit nach Diensten suchen. Für einen gefundenen Dienst kann der Dienstenutzer beim Provider eine XML-basierte Beschreibung abrufen (WebService Description Language, WSDL). Ein daraus generiertes Interface dient dann zum ebenfalls XML-basierten Datenaustausch mit Hilfe eines Kommunikationsprotokolls (Simple Open Access Protocol, SOAP). Der Datenaustausch muss jedoch auch bei Webservices nicht immer auf Basis von XML erfolgen. Insbesondere bei großen Datenmengen ist es ratsam, Referenzparameter wie z.B. einen URI zu übergeben und die eigentlichen Daten in einem effizienteren Format zu übertragen. Auf diesem Weg können auch redundante Datenübertragungen zwischen Webservices vermieden werden.

VO-Mitglieder können auf diese Weise mittels WSRF Archiv-Dienste für Ingest und Access anfordern. Der LZA-Provider kann als Vermittler (broker) zu den anzubindenden Archiven betrachtet werden. Der Benutzer übergibt via WSRF dem Provider die zu archivierenden Objekte, die zugehörigen Metadaten und das Zielarchiv. Der Provider kann danach die Einreichung vorbereiten, via WS mit dem Zielarchiv Kontakt aufnehmen und die Archivobjekte nebst Metadaten übergeben.

Die beschriebene Grid-basierte LZA-Infrastruktur setzt voraus, dass die zu integrierenden Archive ebenfalls Webservices anbieten. Dies ist bei allen von den CGs angegebenen Archiven der Fall. Auch die im Abschnitt 2.2.3 behandelten institutionellen Repositorien unterstützen ebenfalls alle Webservices.

Eine Autorisierung der Benutzer zum Zugriff auf LZA-Dienste kann indirekt über deren Authentifikation mittels VO-Zertifikat erfolgen. Der LZA-Grid-Service-Provider authentifiziert sich stellvertretend für die VO-Mitglieder bei den Archiven.

Das beschriebene Konzept erfüllt alle Anforderungen der CGs, da es einen vereinheitlichten Zugriff auf alle Archivarten aus dem Grid ermöglicht. Die Ingest- und Access-Funktionen

können wie in Abb. 3 dargestellt, auch über ein entsprechend angepasstes Portal (z. B. mit portlets in Gridsphere) genutzt werden. Dies ermöglicht es auch weniger versierten Benutzern, in Grid-Umgebungen auf LZA-Funktionalitäten zuzugreifen.

Im Folgenden werden die einzelnen Archivprozesse ausführlicher beschrieben und die damit verbundenen Synergiepotenziale verdeutlicht.

## **5.1 Ingest-Prozess**

Beim Ingest-Prozess müssen, wie schon erwähnt, vom Benutzer Angaben zum Zielarchiv, die Archivobjekte und zugehörigen Metadaten an den LZA-Provider übergeben werden. Dieser überprüft zunächst die Vollständigkeit der Metadaten und analysiert bzw. validiert die Formate der eingereichten Archivobjekte. Anschließend fragt er beim Zielarchiv via WS nach, ob diese Formate unterstützt werden. Bei Bedarf kann der LZA-Provider notwendige Formatmigrationen anstoßen. Je nach Umfang können diese zur Unterstützung des LZA-Providers auch auf anderen Grid-Ressourcen der VO durchgeführt werden. Wenn die o. g. Formatüberprüfungen erfolgreich waren, übergibt der LZA-Provider den Archivierungsauftrag an das Zielarchiv.

Durch den Ingest über einen LZA-Provider können mehrere Archive über eine einheitliche Schnittstelle angesprochen werden. Außerdem ergibt sich für die Archive eine Entlastung, da unvollständige Eingaben oder nicht archivfähige Formate frühzeitig erkannt werden und dadurch der Ingest-Prozess auf der Archivseite nicht gestartet werden muss.

Es sind jedoch noch weitergehende Entlastungen der Archive vorstellbar. So könnten im Grid vollständige SIPs für die Zielarchive aufbereitet werden, die dann direkt via WS an die Datenmanagement-Komponente übergeben werden. In diesem Fall müssten die Überprüfungen beim Ingest nicht doppelt ausgeführt werden. Außerdem könnten über Indexierungsprozesse auf Grid-Ressourcen zusätzliche beschreibende Metadaten erzeugt oder Ontologiebasierte Analysen der eingereichten Inhalte durchgeführt werden. Analog zur Indexierung von textuellen Inhalten könnten bei Datenarchiven rechenintensive Verarbeitungsschritte auf den Primärdaten (wie z.B. schnelle Fouriertransformationen) im Grid ausgeführt werden.

Ein weiteres interessantes Synergiepotenzial ergibt sich durch die kombinierte Einreichung zusammengehöriger Archivobjekte, die für unterschiedliche Zielarchive bestimmt sind. Nach den von der DFG herausgegebenen „Grundsätze zur Sicherung guter wissenschaftlicher Praxis“ sollten die einer Publikation zugrunde liegenden Daten mindestens 10 Jahre aufbewahrt und öffentlich bereitgestellt werden, damit die publizierten Ergebnisse auch von anderen Forschern überprüfbar sind (vgl. hierzu auch [Klump 2007]). Wenn die Wissenschaftler gleichzeitig Daten und Publikationen einreichen, können wechselseitige Querverweise auf die Objekte leicht eingepflegt werden. Als Querverweise können Persistent Identifier (PI) dienen, die eine vom physischen Speicherort unabhängige Referenz auf Inhalte darstellen. Der PI wird von einer (oder mehreren) zentralen Instanz aufgelöst und liefert eine (oder mehrere) Adressen, meist in Form von URLs (Universal Resource Locator) auf die dort bereitgestellten Inhalte. Der Vorteil von PIs besteht darin, dass die damit referenzierten Objekte auch dann noch gefunden werden, wenn sich die physischen Speicherorte ändern. Ein gutes Beispiel aus der Praxis ist der gemeinsam von der TIB Hannover und dem WDCC entwickelte Prozess zur

Publikation von Primärdaten im STD-DOI-Profil (Scientific Technical Data – Digital Object Identifier).

Die den Daten zugeordneten PIs können bei einer kombinierten Einreichung mit einer Publikation in deren Metadaten aufgenommen werden. Auf diese Weise können automatische Verknüpfungen zwischen zusammengehörenden Archivobjekten erfolgen.

## **5.2 Access-Prozess**

Durch den Access-Prozess sollen anhand von Suchanfragen der Benutzer Inhalte in den Archiven gefunden und bereitgestellt werden. Die Suche erfolgt normalerweise anhand der Metadaten. Bei Dokumenten besteht auch die Möglichkeit, im Volltext zu suchen. Die Volltextsuche ist jedoch sehr zeitaufwändig. Wenn die Anfragen nicht direkt an die Archive, sondern an einen LZA-Provider gerichtet werden, kann dieser – stellvertretend für die Benutzer – via WS alle integrierten Archive ansprechen und die von diesen gefundenen Ergebnisse an den Benutzer weiterleiten. Ähnlich wie bei einer Suchmaschine erhält dieser so einen umfassenden Überblick über die verfügbaren Publikationen und Daten. Gegenüber Einzelabfragen stellt diese Vorgehensweise daher ein weiteres Synergiepotenzial dar, da die Kombination von Grid-Technologien und LZA höherwertige Informationen für die Wissenschaftler liefern.

Um die Suchprozesse zu beschleunigen und die integrierten Archive (und Netzwerkverbindungen) von Suchanfragen zu entlasten, könnte man die Suche alternativ auch ins Grid verlagern. Hierzu müssten in regelmäßigen Zeitabständen Metadaten aus den Archiven in eine Grid-interne Datenbank importiert werden. Da die meisten Archive das standardisierte OAI-PMH unterstützen, kann man hiermit die Metadaten sammeln. Durch die Speicherung in einer Grid-internen Datenbank ergibt sich eine engere Kopplung der Metadaten. Dadurch werden die Suchanfragen beschleunigt und es können komplexere Anfragen in kürzerer Zeit bedient werden. Die erreichbaren Synergien bestehen also im Wesentlichen aus dem Zeitgewinn bei der Suche und der Entlastung von Archiven sowie Netzwerkverbindungen.

Das regelmäßige Abfragen von Metadaten ist nicht sehr effizient, da eigentlich immer nur die Änderungen zum bisherigen Stand in die Grid-interne Datenbank eingespeist werden müssen. Diese inkrementelle Ergänzung könnte über einen WS des LZA-Providers durch die Archive erfolgen. Diese könnten bei jedem erfolgreichen Ingest eines Archivobjekts dem LZA-Provider die dazu gehörigen Metadaten zuschicken.

## **5.3 Preservation-Prozess**

Vor allem in Dokumentarchiven ist eine permanente Technologiebeobachtung unerlässlich, da sich die Dokumentformate häufiger ändern können. Die Technologiebeobachtung wird durch Format Registries wie GDFR (<http://hul.harvard.edu/gdfr/>) und PRONOM (<http://www.nationalarchives.gov.uk/pronom/>) unterstützt. Wenn sich die Archiv-Administration für eine Formatmigration entscheidet, ist meist eine Vielzahl von Objekten betroffen. Es genügt dann aber in der Regel nicht, das Archiv nur nach Dokumenten im alten Format abzusuchen. Häufig sind auch andere Dateien zu migrieren, wenn diese abhängige Objekte im alten Format enthalten. So kann es beispielsweise erforderlich sein, in einem Dokument ein-

gebettete Bildformate herauszufiltern, diese zu migrieren und dann wieder in das ursprüngliche Dokumente einzubauen.

Die Migration vieler Objekte stellt eine große Arbeitslast für ein Archiv dar. Andererseits gibt es im Grid viele Rechenressourcen, die diese Arbeitslast übernehmen können. Der LZA-Provider kann hierzu in seinem Dienstverzeichnis Migrationsdienste als Webservices für die Archive anbieten. Um die Last dynamisch zu verteilen, sollte der LZA-Provider diese Dienste auch wenig ausgelasteten Grid-Ressourcen vermitteln können. Über WS senden die Archive dann die zu migrierenden Dokumente ins Grid und erhalten diese nach der Migration von dort zurück. Die Synergiepotenziale des Preservation-Prozesses bestehen darin, dass die Archive deutlich entlastet und die vorhandenen Grid-Ressourcen besser ausgelastet werden können. Das letztgenannte Synergiepotenzial setzt überschüssige Rechenkapazitäten bei den Grid-Ressourcen voraus. Wie jedoch schon eingangs erwähnt wurde, ist das gegenseitige Geben und Nehmen eine Grundvoraussetzung, um überhaupt Synergien erreichen zu können.

Neben der durch Technologiebeachtung ausgelösten Formatmigration sind auch regelmäßige Validierungen der AIPs durch freie Grid-Ressourcen denkbar. Auch diese Dienstanforderungen werden vom LZA-Provider entgegen genommen und an einen Dienstanbieter auf einer wenig ausgelasteten Grid-Ressource weitergeleitet. Wegen des erheblichen Overheads zur XML-Codierung ist bei großen Datenvolumen eine Binärcodierung oder eine Referenzierung der Daten (z.B. mittels einer URI) in den XML-Nachrichten des WS ratsam. Dabei entfällt der erhebliche und leistungsmindernde Zeitbedarf zum Codieren der Daten in XML bzw. zum Parsen der übertragenen XML-Nachrichten. Die AIP-Container müssen nach der Übertragung beim Dienstanbieter entpackt und die darin enthaltenen Dokumente validiert werden. Werden bei der Überprüfung alte Dokumentformate vorgefunden, so können diese von der Grid-Ressource migriert werden. Nach der Bearbeitung sind alle Dokumentformate wieder auf dem neuesten Technologiestand. Sie werden in einen AIP-Container eingepackt und schließlich in das Archiv zurück übertragen. Da zur Übertragung der AIPs große Netzwerkbandbreiten erforderlich sind, sollte sie möglichst zu Zeiten geringer Netzwerkauslastung ausgeführt werden. Generell stellen die zusätzlich benötigten Netzwerkressourcen einen erheblichen Overhead bei der Nutzung des mit dem Preservation-Prozess verbundenen Synergiepotenzials dar.

## **5.4 Archival Storage**

Im Folgenden werden mögliche Synergiepotenziale bzgl. der Archival Storage-Komponente betrachtet. Der Archival Storage eines typischen Archivs besteht aus Servern, die ein hierarchisches Speichersystem (Hierarchical Storage Management, HSM) aus Festplatten und Magnetbändern realisieren. Die Bändern sind i.a. über robotische Systeme zugreifbar und die Speichereinheiten über speziell für diese Aufgabe optimierte Netzwerke (Storage Area Network, SAN) miteinander verbunden. Auf ein solches HSM-Speichersystem greift die Datenmanagement-Komponente des Archivs über ein Netzwerk zu und verwaltet mit Hilfe einer Datenbank die gespeicherten Inhalte.

Auch diese Komponente kann durch die Grid-Ressourcen unterstützt werden. Das mit Speicherkapazität zu unterstützende Archiv muss hierzu über eine Daten-Grid-Middleware (z.B. SRB oder dCache) mit Speicherressourcen des Grids verbunden werden, welche die gleiche Middleware verwenden. Nun kann das HSM-Speichersystem vom Grid freigegebenen

Speicherplatz einbinden. Die Daten-Grid-Middleware ermöglicht es, dass bestimmte Verzeichnisse auf einem entfernten Grid-Knoten repliziert werden. Damit können Ausfälle von Speicherressourcen toleriert werden. Diese Option bietet gegebenenfalls für die Kooperation mehrerer Archive Synergiepotenziale, da sich dadurch gegenseitige Datensicherungen erstellen lassen. Der gewünschte Grad an Fehlertoleranz kann über die Anzahl der Repliken beliebig skaliert werden, d.h. für Verzeichnisse mit besonders erhaltenswerten Inhalten sollten auch möglichst viele Repliken und weit entfernte Speicherorte vorgesehen werden.

## 6 Handlungsempfehlungen

Wir wollen zunächst alle im letzten Abschnitt identifizierten Synergiepotenziale zwischen LZA und Grid-Technologien noch einmal in einer Übersicht zusammenfassen. Dabei werden sie wie in Abbildung 2 nach den beiden dort verwendeten Kategorien geordnet.

### 1. Grid nutzt LZA

- Nahtlose Integration bestehender und neuer Archive
- Kombinierte Einreichung von Publikationen und zugehörigen Daten
- Bessere Auslastung von Grid-Ressourcen

### 2. LZA nutzt Grid

- Formatvalidierung im Grid ausführen
- Test der Archivfähigkeit bzw. Migration vor Ingest
- Vorbereitung von SIP für Ingest
- Zusätzliche Metadaten durch Indexierung
- Ontologiebasierte Analyse
- Metadatenbasierte Suche im Grid
- Formatmigrationen im Grid
- Migration eingebetteter Formate
- Regelmäßige Validierung und Migration von AIPs
- Archive durch Grid-Speicherressourcen unterstützen

Gegenstand der LZA sind Informationen in digitaler Darstellung, die für die Zukunft erhalten werden sollen. Die nahtlose Integration bestehender und neuer Archive liefert für die Grid-Anwender einen bedeutenden Mehrwert, weil damit ein einheitlicher Zugang zu einer bislang noch nicht verfügbaren Funktionalität im Grid geschaffen wird. Die kombinierte Einreichung von Publikationen und zugehörigen Daten direkt aus der Grid- bzw. eScience-Umgebung heraus bietet den Vorteil, dass konsistente Bezüge zwischen Experimenten bzw. Simulationen und den daraus abgeleiteten Schlussfolgerungen hergestellt werden. Die bessere Auslastung von Grid-Ressourcen führt dazu, dass die Investitionen in die Infrastruktur effizienter genutzt werden und dass folglich auch in Zukunft verstärkt in den Ausbau von Grid-Technologien investiert wird.

Die auf diese Weise freigesetzte Synergie kann für vielfältige rechenintensive Aufgaben im Bereich der LZA genutzt werden und so die Archive entlasten. Beispiele dafür findet man in der obigen Auflistung. Durch die zusätzliche Rechenkapazität des Grids entstehen insbesondere auch weitere Mehrwerte, welche die Qualität der Suchanfragen verbessern (z.B. zusätzli-

che Metadaten durch Indexierung) und neue Methoden der wissensbasierten Suche (z.B. ontologiebasierte Analyse) für die LZA erschließen. Über die Nutzung von Rechenleistung hinaus können die Archive auch von verteilten Speicherressourcen eines Daten-Grids profitieren.

Um die oben genannten Synergiepotenziale freizusetzen, empfehlen wir den Einsatz einer serviceorientierten LZA-Infrastruktur gemäß Abbildung 3, die auf der Basis standardisierter und weitverbreiteter Middleware wie z.B. Globus Toolkit Version 4 implementiert werden kann. Zur Entwicklung der benötigten Funktionalitäten sollen Webservices eingesetzt werden, die als selbstbeschreibende Softwarekomponenten eine dynamische Ausführung von verteilt bereitgestellten Diensten ermöglichen.

Es ist nicht ratsam, die Webservices für die jeweiligen Synergiepotenziale losgelöst von einander zu entwerfen. Stattdessen sollte eine auf die gesamte servicebasierte LZA-Infrastruktur abgestimmte Suite von Webservices mit orthogonalem Funktionsumfang entwickelt werden. Orthogonalität bedeutet, dass die einzelnen Dienste möglichst generisch und wenig miteinander korreliert sind, um durch ihre Kombination weitere höherwertige Webservices erzeugen zu können. Bei der Konzipierung der einzelnen Webservices ist also darauf zu achten, dass einerseits möglichst universelle und leicht kombinierbare Module entstehen, andererseits aber dadurch der Kommunikationsaufwand nicht unnötig in die Höhe getrieben wird. Vor allem Übertragungen von großen Datenmengen erzeugen wegen der XML-Codierung einen immensen Overhead. Das Problem besteht also besonders bei Daten oder Digitalisaten.

Die von einem LZA-Grid-Service-Provider bereitgestellten Webservices sollten außerdem konfigurierbar sein. Auf diese Weise sind flexible Anpassungen an die jeweiligen Erfordernisse innerhalb einer eScience-Umgebung möglich. Ein LZA-Grid-Service-Provider könnte auch Webservices enthalten, die eine Kopplung mit LZA-Infrastrukturen anderer Communities erlauben. Hierdurch wäre eine interdisziplinäre Wissensvernetzung realisierbar.

Um die Archive über Webservices ansprechen zu können, muss eine Spezifikation der Funktionen und Parameter erarbeitet werden, die zur Kommunikation mit dem LZA-Grid-Service-Provider benötigt werden. Es sollte langfristig darauf hin gearbeitet werden, aus einer solchen Spezifikation einen Standard abzuleiten. Archive mit standardisiertem Webservice-Funktionsumfang könnten dann sehr viel leichter in Grid-Umgebungen mit einer LZA-Infrastruktur integriert werden. Weitere Aspekte der Standardisierung werden in einer gesonderten Expertise ausführlich behandelt [Borghoff et al., 2008].

Neben der Entwicklung einer Suite von Webservices müssen auch Methoden zum Scheduling erarbeitet werden. Da sowohl eine gute Auslastung der Ressourcen als auch eine schnelle Reaktionen der Webservices wünschenswert sind, muss durch einen Scheduler für eine optimierte dynamische Zuordnung der auszuführenden Webservices zu den Grid-Ressourcen gesorgt werden. Diese Komponente hat entscheidenden Einfluss auf die Performanz und damit auch auf die Akzeptanz einer LZA-Infrastruktur, obwohl sie keineswegs LZA-spezifisch ist, sondern durch die Nutzung von Grid-Technologien begründet wird. Daher empfehlen wir, dass der Lösung des Scheduling-Problems frühzeitig eine besondere Bedeutung zugemessen wird und dass unter Hinzuziehung von Experten aus dem Scheduling-Umfeld geeignete Lösungen erarbeitet werden.

Wir empfehlen weiter, ein Testbed aufzubauen in dem die entwickelten Komponenten evaluiert werden können. Hierzu sollte eine LZA-VO etabliert werden, in die exemplarisch je-

weils ein Vertreter der drei o. g. Archive integriert werden könnte. Als typisches Beispiel für ein Archiv einer Gedächtnisorganisation würde sich kopal anbieten. Es bietet über das frei verfügbare koLibRI-Tool bereits eine gut dokumentierte und anpassbare Webservice-Schnittstelle. Auch die Datenarchive des IVOA oder des WDCC unterstützen Zugänge über Webservices und kommen daher für die Integration in einem Testbed in Frage. Als Beispiele für institutionelle Repositorien könnten dSpace oder Fedora benutzt werden. Beide verfügen über konfigurierbare und dokumentierte Webservice-Schnittstellen.

Mit Hilfe des institutionellen Repositories sollte auch die Unterstützung eines Archivs durch Speicherressourcen aus dem Grid erprobt werden. Wir empfehlen hierzu, das Repository auf einer Grid-Ressource (Server) der LZA-VO zu installieren und an ein dort ebenfalls vorhandenes Daten-Grid (z.B. SRB oder dCache) anzubinden. Hiermit könnte dann die Stabilität und Performanz eines quasi vollständig gridifizierten Archivs analysiert werden.

Anhand des Testbeds soll die Performanz und Akzeptanz der serviceorientierten LZA-Architektur evaluiert und optimiert werden. Erst danach empfehlen wir einen Piloteinsatz in ausgewählten CGs und nach einem weiteren Optimierungszyklus die Aufnahme des Produktionsbetriebs. Darauf aufbauend könnten später auch verteilte CG-spezifische Arbeitsumgebungen realisiert werden, die zusätzliche Werkzeuge zur Kollaboration und zum Wissensmanagement bereitstellen.

## **Danksagung**

Wir bedanken uns bei allen, die durch konstruktive Kritik halfen, die Qualität der Expertise zu verbessern. Außerdem danken wir den Vertretern der drei CGs, des WDCC und des WDC-RSAT für die Beantwortung unserer Fragen.

## 7 Literatur

Allcock et al.: *GridFTP: Protocol Extensions to FTP for the Grid*. Internet-Draft der IETF, 2001

Antonioletti et al.: *OGSA-DAI 3.0 - The Whats and the Whys*. Proceedings of the UK e-Science All Hands Meeting 2007, September 2007

Baker et al.: *A Fresh Look at the Reliability of Long-term Digital Storage*. Proceedings of EuroSys, April 2006.

Borghoff et al.: *Vergleich bestehender Archivierungssysteme*. nestor-Expertise, 2005

Borghoff et al.: *Standards und Standardisierung im Kontext von Grid-Technologien und Langzeitarchivierung*. nestor-Expertise, 2008

Bos et al.: *From Shared Databases to Communities of Practice: A Taxonomie of Collaboratories*. Journal of Computer-Mediated Communication, 12(2), Article 16, 2007

Braun et al.: *Im Wissensnetz: Linked Information Processes in Research Networks*. GES2007, <http://www.ges2007.de>, 2007

D-Grid AK5: *Daten- und Informationsmanagement*. Zusammenfassendes Abschlusspapier des D-Grid AK5, 2004

Enke et al.: *AstroGrid-D: Enhancing Astronomic Science with Grid Technology*. GES2007, <http://www.ges2007.de>, 2007

eSciDoc: *Das Projekt eSciDoc*. <http://www.escidoc-project.de/de/das-project-escidoc.html>

Ferreira et al.: *Introduction to Grid Computing with Globus*. IBM Redbooks, 2005

Foster I. und Kesselman C.: *Computational Grids*. In: The Grid: Blueprint for a Future Computing Infrastructure, I. Foster und C. Kesselman (Eds.), Morgan Kaufmann Publishers, 1998

Foster I.: *The Anatomy of the Grid: Enabling Scalable Virtual Organizations*. In Proceedings of the 7th international Euro-Par Conference Manchester on Parallel Processing (August 28 - 31, 2001), Lecture Notes in Computer Science, vol. 2150, Springer-Verlag, 2001

Foster I.: *Globus Toolkit Version 4: Software for Service-Oriented Systems*. In NPC 2005, LNCS 3779, IFIP International Federation for Information Processing, 2005

Foster I.: *Globus Toolkit Version 4: Software for Service-Oriented Systems*. Journal of Computer Science and Technology, Vol. 21, No 4., 2006

Fuhrmann P. und Gülzow V.: *dCache, storage system for the future*. In: Proceedings of the 12th International Euro-Par Conference, LNCS 4128, S. 1106 – 1113, Springer, 2006

Fuhrmann et al.: *dCache, scalable managed storage*. GES2007, <http://www.ges2007.de>, 2007

Gietz et al.: *TextGrid and eHumanities*. In: E-SCIENCE '06: Proceedings of the Second IEEE International Conference on e-Science and Grid Computing, S. 133 ff., IEEE Computer Society, 2006

GridFTP: *Universal Data Transfer for the Grid*. White Paper, The Globus Project, 2000

Hefke et al.: Präsentation des Projektes auf dem Kick-Off der BMBF-Förderinitiative eScience-Wissensvernetzung in Bonn, 06.04.06

Helmes L. und Razum M.: *eSciDoc – Stand und Ausblick*. FIZ Karlsruhe, Oktober 2006

Hiller W. und Fritsch B.: *The Collaborative Climate Community Data and Processing Grid (C3-Grid) – Architecture and Implementation Issues*. GES2007, <http://www.ges2007.de>, 2007

IBM: *Digital Information Archiving System (DIAS)*, <http://www-05.ibm.com/nl/dias/>, 2008

Klump J.: *Anforderungen von eScience und Grid-Technologie an die Archivierung wissenschaftlicher Daten*. nestor-Expertise, 2007

Lautenschlager M. und Stahl W.: *Long-Term Archiving of Climate Model Data at WDC Climate and DKRZ*. In: E. Mikusch (Ed.): PV2007 – Ensuring the Long-Term Preservation and Value Adding to Scientific and Technical Data, Conference Proceedings. DLR, German Remote Sensing Data Center, Oberpfaffenhofen, 2007

Meyers Lexikonverlag: *Meyers Lexikon Online 2.0*. Seitenname: Synergie, Herausgeber: Bibliographisches Institut & F. A. Brockhaus AG, <http://lexikon.meyers.de/index.php?title=Synergie&oldid=131410>, 2007

Moore et al.: *Data Grids, Digital Libraries and Persistent Archives: An Integrated Approach to Publishing, Sharing and Archiving Data*. In: *Proceedings of the IEEE*, Vol. 93, No. 3, 2005

Moore et al.: *Production Storage Resource Broker Data Grids*. In: Proceedings of the Second IEEE International Conference on e-Science and Grid Computing (e-Science'06), IEEE Computer Society, 2006

NIRVANA: *Data Sharing with Storage Resource Broker – Enabling Collaboration in Complex Distributed Environments*. White Paper, Nirvana division <http://www.nirvanastorage.com/products/whitepapers.htm>, 2007

OAIS: *Reference Model for an Open Archival Information System (OAIS)*. CCSDS 650.0-B-1 Blue Book, Washington, DC, CCSDS, [www.classic.ccsds.org/documents/pdf/CCSDS-650.0-B-1.pdf](http://www.classic.ccsds.org/documents/pdf/CCSDS-650.0-B-1.pdf), 2002.

ONTOVERSE : *Projektbeschreibung zu ONTOVERSE*. <http://www.ontoverse.org/>

Quinn et al.: *The Management, Storage and Utilization of Astronomical Data in the 21<sup>st</sup> Century, A Discussion Paper for the OECD Global Science Forum*. <http://www.ivoa.net/pub/info/OECD-QLH-Final.pdf>, 2004

Rajasekar et al.: *A Prototype Rule-based Distributed Data Management System*. HPDC workshop on Next Generation Distributed Data Management, May 2006, Paris, France, 2006

Schiffmann W.: *Technische Informatik, Bd. 2, Grundlagen der Computertechnik*. 5. Auflage, Springer-Verlag, Heidelberg, Berlin, 2005

O'Shields M.E und Lunsford II, P.J.: *WebDAV: A Web-Writing Protocol and More*. In: Journal of Industrial Technology, Vol. 20, No.2, NAIT, 2004

Steinmetz R. und Wehrle K.: *Peer-to-Peer-Networking & -Computing*. In: Informatik Spektrum, Vol. 20, No. 1, Springer, 2004

Stockinger H.: *Defining the grid: a snapshot on the current view*. The Journal of Supercomputing, Vol. 42, No. 1, Springer, 2007

Ulbrich et al.: *C3-Grid – ein Werkzeug für die Klimaforschung*. DACH Meteorologentagung, [http://meetings.copernicus.org/dach2007/download/DACH2007\\_A\\_00354.pdf](http://meetings.copernicus.org/dach2007/download/DACH2007_A_00354.pdf), 2007

Venugopal et al.: *A Taxonomy of Data Grids for Distributed Data Sharing, Management, and Processing*. ACM Computing Surveys, Vol. 38, March 2006

WIKINGER: *Projektbeschreibung zu WIKINGER*. [http://www.wikinger-escience.de/xwiki/bin/download/Main/WebHome/WIKINGER\\_20060601.pdf](http://www.wikinger-escience.de/xwiki/bin/download/Main/WebHome/WIKINGER_20060601.pdf), 2006

## 8 Anhang

Neben den explizit im Rahmen dieser Studie genannten Grid-, eScience-, und dLZA-Projekten wurden zahlreiche weitere Projekte betrachtet, deren Ideen, Konzepte und Ergebnisse in Summe die Identifikation der Synergiepotenziale und den Entwurf der vorgeschlagenen serviceorientierten LZA-Infrastruktur beeinflusst haben. Die nachfolgende Übersicht bietet eine Aufstellung aller dieser Projekte, die anhand ihrer geographischen Herkunft in die Kategorien National, Europäisch und International eingeteilt werden.

### 8.1 Deutsche Projekte

Projektname	URL
D-Grid	<a href="http://www.d-grid.de">www.d-grid.de</a>
DGI	<a href="http://Dgi.d-grid.de">Dgi.d-grid.de</a>
AstroGrid-D	<a href="http://www.gac-grid.de">www.gac-grid.de</a>
C3-Grid	<a href="http://www.c3grid.de">www.c3grid.de</a>
HEP Community Grid	<a href="http://www.d-grid.de/index.php?id=44">www.d-grid.de/index.php?id=44</a>
InGrid	<a href="http://www.ingrid-info.de">www.ingrid-info.de</a>
MediGRID	<a href="http://www.medigrid.de">www.medigrid.de</a>
TextGrid	<a href="http://www.d-grid.de/index.php?id=167">www.d-grid.de/index.php?id=167</a>
Kompetenznetzwerk LZA	<a href="http://langzeitarchivierung.de">langzeitarchivierung.de</a>
BABS	<a href="http://www.babs-muenchen.de">www.babs-muenchen.de</a>
Kopal	<a href="http://kopal.langzeitarchivierung.de">kopal.langzeitarchivierung.de</a>
MyCoRe	<a href="http://www.mycore.de">http://www.mycore.de</a>
eSciDoc	<a href="http://www.escidoc-project.de">www.escidoc-project.de</a>
FRESCO	<a href="http://www.ipsi.fraunhofer.de/i-info/de/content/view/97/0/">http://www.ipsi.fraunhofer.de/i-info/de/content/view/97/0/</a>
STEMNET	<a href="http://www.dl-forum.de/deutsch/projekte/projekte_2681_DEU_HTML.htm">http://www.dl-forum.de/deutsch/projekte/projekte_2681_DEU_HTML.htm</a>
SYNERGIE	<a href="http://www.dl-forum.de/deutsch/projekte/projekte_2683_DEU_HTML.htm">http://www.dl-forum.de/deutsch/projekte/projekte_2683_DEU_HTML.htm</a>
HyperImage	<a href="http://www.dl-forum.de/deutsch/projekte/projekte_2688_DEU_HTML.htm">http://www.dl-forum.de/deutsch/projekte/projekte_2688_DEU_HTML.htm</a>
Im Wissensnetz	<a href="http://www.dl-forum.de/deutsch/projekte/projekte_2538_DEU_HTML.htm">http://www.dl-forum.de/deutsch/projekte/projekte_2538_DEU_HTML.htm</a>
WIKINGER	<a href="http://www.imk.fraunhofer.de/de/wikinger">http://www.imk.fraunhofer.de/de/wikinger</a>
WISENT	<a href="http://www.dl-forum.de/deutsch/projekte/projekte_2370_DEU_HTML.htm">http://www.dl-forum.de/deutsch/projekte/projekte_2370_DEU_HTML.htm</a>
Ontoverse	<a href="http://www.dl-forum.de/deutsch/projekte/projekte_2381_DEU_HTML.htm">http://www.dl-forum.de/deutsch/projekte/projekte_2381_DEU_HTML.htm</a>
links4science	<a href="http://www.dl-forum.de/deutsch/projekte/projekte_2764_DEU_HTML.htm">http://www.dl-forum.de/deutsch/projekte/projekte_2764_DEU_HTML.htm</a>

## 8.2 Europäische Projekte

Projektname	URL
EPoSS	<a href="http://www.smart-systems-integration.org">www.smart-systems-integration.org</a>
ASP-BP	<a href="http://www.asp-bp.org">www.asp-bp.org</a>
BIOGRID	<a href="http://www.bio-grid.net">www.bio-grid.net</a>
EGSO	<a href="http://www.egso.org">www.egso.org</a>
COG	<a href="http://www.cogproject.org">www.cogproject.org</a>
DATAGRID	<a href="http://www.edg.org">www.edg.org</a>
DAMIEN	<a href="http://www.hlrs.de/organization/pds/projects/damien">www.hlrs.de/organization/pds/projects/damien</a>
DATATAG	<a href="http://datatag.web.cern.ch/datatag">datatag.web.cern.ch/datatag</a>
FLOWGRID	<a href="http://www.unizar.es/flowgrid">www.unizar.es/flowgrid</a>
EUROGRID	<a href="http://www.eurogrid.org">www.eurogrid.org</a>
CROSSGRID	<a href="http://www.eu-crossgrid.org">www.eu-crossgrid.org</a>
GEMSS	<a href="http://www.ccll.nece.de/gemss">www.ccll.nece.de/gemss</a>
GRACE	<a href="http://www.eu-grasp.net">www.eu-grasp.net</a>
GRASP	<a href="http://www.eu-grasp.net">www.eu-grasp.net</a>
GRIA	<a href="http://www.gria.org">www.gria.org</a>
GRIDLAB	<a href="http://www.gridlab.org">www.gridlab.org</a>
GRIDSTART	<a href="http://www.gridstart.org">www.gridstart.org</a>
GRIP	<a href="http://www.grid-interoperability.org">www.grid-interoperability.org</a>
MAMMOGRID	<a href="http://Mammogrid.vitamib.com">Mammogrid.vitamib.com</a>
MOSES	<a href="http://www.hum.ku.dk/moses">www.hum.ku.dk/moses</a>
OPENMOLGRID	<a href="http://www.openmolgrid.org">www.openmolgrid.org</a>
SELENE	<a href="http://www.dcs.bbk.ac.uk/selene">www.dcs.bbk.ac.uk/selene</a>
WEBSI	<a href="http://www.ib-ia.com/websi">www.ib-ia.com/websi</a>
DataMiningGrid	<a href="http://www.datamininggrid.org">www.datamininggrid.org</a>
Trustcom	<a href="http://www.eu-trustcom.com">www.eu-trustcom.com</a>
ICEAGE	<a href="http://www.iceage-eu.org">www.iceage-eu.org</a>
Gridcoord	<a href="http://www.gridcoord.org">www.gridcoord.org</a>
OMII-Europe	<a href="http://www.omii-europe.com">www.omii-europe.com</a>
SIMDAT	<a href="http://www.simdat.eu">www.simdat.eu</a>
CoreGRID	<a href="http://www.coregrid.net">www.coregrid.net</a>
ASG	<a href="http://www.asg-platform.org">www.asg-platform.org</a>
Deisa	<a href="http://www.deisa.org">www.deisa.org</a>
DBE	<a href="http://www.digital-ecosystem.org">www.digital-ecosystem.org</a>
DIP	<a href="http://dip.semanticweb.org">dip.semanticweb.org</a>
K-Wf Grid Experience	<a href="http://www.kwfgrid.eu">www.kwfgrid.eu</a>
eLegi	<a href="http://www.elegi.org">www.elegi.org</a>
EGEE II	<a href="http://www.eu-egee.org">www.eu-egee.org</a>
InteliGrid	<a href="http://www.inteliGrid.com">www.inteliGrid.com</a>
aKogrimo	<a href="http://www.mobilegrids.org">www.mobilegrids.org</a>
NextGRID	<a href="http://www.nextgrid.org">www.nextgrid.org</a>
ONTOGRID	<a href="http://www.ontogrid.net">www.ontogrid.net</a>
Provenance	<a href="http://twiki.gridprovenance.org/bin/view/Provenance/WebHome">twiki.gridprovenance.org/bin/view/Provenance/WebHome</a>
UniGrids	<a href="http://www.unigrids.org">www.unigrids.org</a>
AOLA	<a href="http://www.ifs.tuwien.at/~aola/">www.ifs.tuwien.at/~aola/</a>
archiPOL	<a href="http://www.archipol.nl">www.archipol.nl</a>
EPrints	<a href="http://www.eprints.org/">http://www.eprints.org/</a>

### **8.3 Internationale Projekte**

Projektname	URL
fedora	<a href="http://www.fedora.info/">www.fedora.info/</a>
dSpace	<a href="http://www.dspace.org/">www.dspace.org/</a>
PADI	<a href="http://www.nla.gov.au/padi/">www.nla.gov.au/padi/</a>
LOCKSS	<a href="http://www.lockss.org">www.lockss.org</a>